AD NUMBER

ADB246269

NEW LIMITATION CHANGE

TO

Approved for public release, distribution unlimited

FROM

Distribution authorized to U.S. Gov't. agencies only; Proprietary Info.; Oct 98. Other requests shall be referred to U.S. Army Medical Research and Materiel Command, Fort Detrick, MD 21702-5012.

AUTHORITY

USAMRMC ltr dtd 4 Jan 2000

AD_____

GRANT NUMBER DAMD17-94-J-4353

TITLE: Regulation of Estrogen Receptor Transcription in Breast Carcinoma

PRINCIPAL INVESTIGATOR: Ronald J. Weigel, M.D., Ph.D.

CONTRACTING ORGANIZATION: Stanford University
Stanford, California 94305

REPORT DATE: October 1998

TYPE OF REPORT: Final

PREPARED FOR: Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, Maryland 21702-5012

NOTICE

USING GOVERNMENT DRAWINGS, SPECIFICATIONS, OR OTHER DATA INCLUDED IN THIS DOCUMENT FOR ANY PURPOSE OTHER THAN GOVERNMENT PROCUREMENT DOES NOT IN ANY WAY OBLIGATE THE U.S. GOVERNMENT. THE FACT THAT THE GOVERNMENT FORMULATED OR SUPPLIED THE DRAWINGS, SPECIFICATIONS, OR OTHER DATA DOES NOT LICENSE THE HOLDER OR ANY OTHER PERSON OR CORPORATION; OR CONVEY ANY RIGHTS OR PERMISSION TO MANUFACTURE, USE, OR SELL ANY PATENTED INVENTION THAT MAY RELATE TO THEM.

LIMITED RIGHTS LEGEND

Award Number: DAMD17-94-J-4353
Organization: Stanford University
Location of Limited Rights Data (Pages):

Those portions of the technical data contained in this report marked as limited rights data shall not, without the written permission of the above contractor, be (a) released or disclosed outside the government, (b) used by the Government for manufacture or, in the case of computer software documentation, for preparing the same or similar computer software, or (c) used by a party other than the Government, except that the Government may release or disclose technical data to persons outside the Government, or permit the use of technical data by such persons, if (i) such release, disclosure, or use is necessary for emergency repair or overhaul or (ii) is a release or disclosure of technical data (other than detailed manufacturing or process data) to, or use of such data by, a foreign government that is in the interest of the Government and is required for evaluational or informational purposes, provided in either case that such release, disclosure or use is made subject to a prohibition that the person to whom the data is released or disclosed may not further use, release or disclose such data, and the contractor or subcontractor or subcontractor asserting the restriction is notified of such release, disclosure or use. This legend, together with the indications of the portions of this data which are subject to such limitations, shall be included on any reproduction hereof which includes any part of the portions subject to such limitations.

THIS TECHNICAL REPORT HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION.

Patricia Maddox _____ 7/16/99 _____

_____ _____

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	October 1998	Final (1 Oct 94 - 30 Sep 98)

4. TITLE AND SUBTITLE
Regulation of Estrogen Receptor Transcription in Breast Carcinoma

5. FUNDING NUMBERS
DAMD17-94-J-4353

6. AUTHOR(S)
Ronald J. Weigel, M.D., Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Stanford University
Stanford, California 94305

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, Maryland 21702-5012

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Distribution authorized to U.S. Government agencies only (proprietary information, Oct 98). Other requests for this document shall be referred to U.S. Army Medical Research and Materiel Command, 504 Scott Street, Fort Detrick, Maryland 21702-5012.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200*

In this study, we performed a functional mapping of the Estrogen Receptor (ER) promoter region and found a region in the 5' untranslated leader sequence that controlled ER transcription in ER-positive breast carcinoma cell lines. This region contained two binding sites for a transcription factor, Estrogen Receptor Factor-1 (ERF-1). ERF-1 was identified by gel shift assay and was present in ER-positive breast and endometrial carcinoma cells, but absent in ER-negative cell lines, thus inferring a role for ERF-1 in the regulation of ER transcription in breast carcinoma. Mutational analysis of the ERF-1 binding site showed that ERF-1 binding was correlated with transcriptional activity. ERF-1 was purified from MCF7 breast carcinoma cells using ion-exchange and DNA affinity chromatography and the cDNA was isolated from a MCF7 expression cDNA library. ERF-1 was identified as AP2γ, a member of the AP2 transcription factor family. When binding specificity of ERF-1 was compared to AP2α using PCR-Assisted Binding Site selection and competitive gel shift assay, it was found that the two proteins recognized the same DNA sequence. Cotransfection of AP2α or AP2γ with an ER promoter reporter construct demonstrated that activation of the ER promoter is dependent upon cell type.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Breast Cancer	80
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Limited

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material.

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

RW ✓ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

RW ✓ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

RW ✓ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

RW ✓ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

Ronald Wenzel 10.27.88

PI - Signature Date

TABLE OF CONTENTS

INTRODUCTION

Background

The expression of estrogen receptor (ER) is intimately associated with the biology of breast carcinoma. Breast carcinomas occurring in postmenopausal women are often ER-positive [1] and many of these tumors express significantly more receptor than normal mammary epithelium [2]. ER-negative breast carcinomas are more likely to occur in young women and these tumors carry a worse prognosis than carcinomas which express ER [3, 4]. Several studies have focused on the function of ER in an attempt to explain the association between ER expression and tumor biology. Mutations have been described in the ER gene of some breast carcinomas that render these altered ER proteins incapable of binding estrogen response elements (ERE) [5, 6] and able to inhibit wild-type ER function [7, 8]. Other studies, however, have found ER mutations which result in a constitutively active receptor which has also been postulated as important to the development of hormone-independent growth [9, 10]. If ER function is influencing the oncogenic process, it is difficult to conceptualize within a single model of oncogenesis the occurrence of mutations which inhibit ER function and mutations which result in constitutive activity. An alternative hypothesis is that mechanisms regulating transcription of the ER gene influence the phenotype of breast carcinoma; within this model, ER-negative carcinomas which do not transcribe the ER gene define a subset of tumors with a more aggressive phenotype. This theory is supported by recent studies which have identified breast carcinoma cell lines that fail to transcribe an apparently normal ER gene [11]. It is therefore possible that defining molecular mechanisms controlling transcription of the ER gene may provide new insight into the biology of breast carcinoma.

Transcription of ER occurs from a number of different promoters, with at least 4 different transcripts identified to date [12, 13]. P1 is the major ER transcriptional start site [14] and is predominantly utilized in human mammary epithelial cells (HMEC) and is the major start site in ER-positive human breast carcinomas [15]. Multiple cap sites have been

identified for the P0 promoter. Studies of the murine ER gene identified 10 cap sites spanning approximately 60 bases [16] and a start site at -1994 (from the P1 cap site) was identified in human cells which would agree closely with the major murine P0 cap site [17]. Transcription from the P0 promoter is characteristic of human endometrial tissue and can account for 12 to 33% of ER transcription in breast carcinoma cells [15]. Recently, several novel transcription start sites have been identified, including two, P_E and P_H, which splice into +163 of the untranslated leader sequence as does the P0 transcript [13]. In transcript P_E, the sequence upstream from the +163 splice site corresponds to -359 to -169 (relative to the P1 start site), a region previously described as part of the intron between P0 and P1. The P_H cap site has been shown to be at least 20 kb upstream of the P1 promoter, suggesting the existence of an alternate promoter for ER.

Purpose of Present Work

The experiments described herein address the technical objectives 1,2,3, and 4 of the U.S. Army Research Grant DAMD17-94-J-4353. In this study, we performed a functional mapping of the ER promoter region from 3.5 kb upstream of the P1 start site to the translation start site and found a region in the 5' untranslated leader sequence that controlled ER transcription in ER-positive breast carcinoma cell lines. This region contained two binding sites for a transcription factor, Estrogen Receptor Factor-1 (ERF-1). ERF-1 was identified by gel shift assay and was present in ER-positive breast and endometrial carcinoma cells, but absent in ER-negative cell lines, thus inferring a role for ERF-1 in the regulation of ER transcription in breast carcinoma. Mutational analysis of the ERF-1 binding site showed that ERF-1 binding was correlated with transcriptional activity. ERF-1 was purified from MCF7 breast carcinoma cells using ion-exchange and DNA affinity chromatography and the cDNA was isolated from a MCF7 expression cDNA library. ERF-1 was identified as AP2γ, a member of the AP2 transcription factor family. When binding specificity of AP2γ was compared to AP2α using PCR-Assisted Binding Site

selection and competitive gel shift assay, it was found that the two proteins recognized the same DNA sequence. Cotransfection of AP2α or AP2γ with an ER promoter reporter construct demonstrated that activation of the ER promoter is dependent upon cell type.

MATERIALS AND METHODS

Cell lines

HMEC were obtained from reduction mammoplasties and maintained in DFCI-1 growth medium as described previously[18]. The ECC-1 ER-positive human endometrial carcinoma cell line was obtained from P. G. Satyaswaroop, Hershey, PA. All other cell lines were obtained from American Type culture Collection, Rockville, MD. Cells were maintained in minimal essential medium (Gibco BRL, Gaithersburg, MD) supplemented with 10% fetal bovine sera (Hyclone, Logan, UT), 25 mM 4-(2-hydroxyethey)-1piperazineethanesulfonic acid (Hepes), 26 mM sodium bicarbonate, 5,000 units/ml penicillin G (Gibco BRL), 5,000 ug/ml streptomycin (Gibco BRL), and 6 ng/ml bovine insulin (Sigma Chemical Company, St. Louis, MO). Cells were incubated at 37°C in 5% CO_2.

Plasmid construction and transfection

The region of DNA in the 5' end of the ER gene was isolated from a human genomic lambda library (Stratagene, La Jolla, CA). The 5' end was sequenced, and a set of primers was constructed, all of which contained a *Bgl*II site for use as upstream primers. All 3' oligonucleotides contained a *Hind*III site. These oligonucleotide primers were used in PCR with cloned DNA as the template. Primers used to generate 5' deletion constructs were ER724 (TACAGATCTGTGGTCCAACATAAACACA), ER586 (TGCAGATCTTCCTAT ATGTATACCC), ER464(CATTAGATCTGCCCTATCTCGGTTACAGTGT), ER375(G GGGAGATCTAACAGAAAGAGAGACAA), ER289 (CCCTAGATCTGTCTTTCGCGT TTAT), ER122 (GGGAGATCTGCCTGGAGTGATG TTTAAG), ER40 (TATGAGATC

TGGAGACCAGTACTTAAAG), and ER0 (CCCAGATCTGGCGGAGGGCGTTCG). Primers used to generate specific 3' ends were +230 (CATAAGCTTGGTCCGTGGCCG CGGGCAGGGT), +210 (CGGGAAGCTTGCAGACCGTGTCCCCGCAGG), +135 (G CCCAAGCTTAGAGGCGACGCAGCGCA), and +0 (CGCCAAGCTTCCTGGGCTCC CGGGCCTC).

PCR products were then subcloned into the *Bgl*II-*Hind*III sites of the luciferase reporter plasmid pGL2-Basic (Promega, Madison, Wis.). To construct the plasmids with 3,500 bp upstream of P1, the 5' region was first subcloned into pBluescript from the genomic lambda clone as a *Sal*I-*Not*I fragment, using the *Sal*I site in lambda DASH and the *Not*I site in the first exon of ER. The ER724-xLUC constructs were each digested with *Xho*I-*Nde*I, removing the 5' region of the ER gene. This region was then replaced with the larger 5' flanking region as an *Xho*I-*Nde*I fragment from the ER gene cloned in pBluescript. In this construct, the *Xho*I site 6 bp upstream of the *Sal*I site in pBluescript was used and was inserted into the *Xho*I site 4 bp upstream of the *Bgl*II site in pGL2-Basic. Each of these plasmids then contained 3,500 bp upstream of P1.

Plasmid DNA was prepared by alkaline lysis and purified by using polyethylene glycol as described previously [19]. All cells were transfected by the calcium phosphate precipitation procedure [20] In these experiments, 30 μg of cloned plasmid DNA was used in transfections of 100-mm-diameter plates with cells at 50 to 60% confluence. In each transfection, 2 μg of a Rous sarcoma virus-driven chloramphenicol acetyltransferase (CAT) expression vector was cotransfected and cells were assayed for luciferase and CAT expression 48 h after transfection. Values presented are luciferase units as measured on a luminometer corrected for transfection efficiency, as determined by CAT assay. Similar results were obained when 15 μg of plasmid DNA was used in transfections.

Preparation of Whole Cell Extracts

Cells were collected by trypsinization. Cell pellets were washed with 1 x PBS and then resuspended at 10^8 cells/ml in microextraction buffer [450 mM NaCl, 20 mM HEPES, pH 7.7, 25% glycerol, 0.2 mM EDTA, 0.5 mM DTT, and proteinase inhibitors]. Cells were sonicated and cellular debris was pelleted at 14,000 rpm in a microfuge. Protein concentration of the supernatant was determined using a Bio-Rad protein assay (Hercules, CA) and ran in the range of 5-15 mg/ml. Extracts were stored at -80°C until use.

Gel Shift Assay

Gel shift assays were performed using 15 ug nuclear cell extract, purified chromatographic fractions or *in vitro* translated protein. Protein was incubated in 1x binding buffer (40 mM KCl/ 20 mM Hepes, pH 7.9/ 1 mM $MgCl_2$/ 0.08 mM ethyleneglycol-bis-(ß-aminoethyl)-N, N, N', N' tetraacetic acid (EGTA)/ 0.4 mM DTT), 4% Ficoll, 12 ug/ml poly dI-dC (Pharmacia), 1 mg/ml Bovine Serum Albumin (BSA) (Sigma) and 0.2 ng of radiolabelled double-stranded oligonucleotide probe in a volume of 25 ul at room temperature for 1 hour. Reactions were electrophoresed on a 4% acrylamide in 0.25X TBE at 4°C. Gels were dried and autoradiographed O/N. In competitive binding assays, unlabeled mutant oligonucleotides were added at 1,000 fold excess. Supershift assays were performed by the addition of 2 ug of affinity-purified rabbit polyclonal antibody directed against human AP2 (Santa Cruz Biochemicals) or with 2 ul of polyclonal La/SS-B antisera obtained as a gift from Michael Bachmann (Mainz, Germany). Probe was prepared by end-labeling the wild-type 30-bp double-stranded ERF-1 binding site with [γ-^{32}P]ATP (3,000 Ci/mmol) using T4 polynucleotide kinase, followed by gel purification on 15% polyacrylamide. Mutant oligonucleotides used as competitors were synthesized to contain specific point mutations in relation to the wild-type ERF-1 sequence.

Purification of ERF-1

MCF7 nuclei were prepared as previously described[21]. Nuclei were resuspended in 0.6 times packed nuclear volume (PNV) of buffer D (10 mM Tris, pH 7.9/100 mM KCl/2 mM $MgCl_2$/0.1 mM EDTA/1 mM DTT/ 1 mM NaMetaBis/0.2 mM phenylmethylsulfonyl fluoride) and Dounce homogenized on ice with 10 strokes. Homogenized nuclei were mixed with 0.06X PNV of 4 M ammonium sulfate (pH 7.9) and extracted with gentle mixing at 4°C for 1 h. Extracted nuclei were pelleted by centrifugation at 25,000 x *g*, 4°C for 20 min and the supernatant was dialyzed for 5.5 h against 5 liters of D-100 buffer (20 mM Hepes, pH 7.9/20% glycerol/ 100 mM KCl/2 mM $MgCl_2$/0.2 mM EDTA/1 mM DTT/ 1 mM NaMetaBis/0.2 mM phenylmethylsulfonyl fluoride). After dialysis, the extract was centrifuged for 20 min at 16,000 x *g*, 4°C. Protein concentration of nuclear extract was determined using a Bio-Rad protein assay system (Bio-Rad) and was generally in the range of 7-15 mg/ml.

ERF-1 activity was monitored throughout the purification by gel shift assay. MCF7 nuclear extract (375 mg) was applied to a 30 ml Q Sepharose Fast Flow (Pharmacia) anion exchange column at 100 mM KCl and eluted with 1 volume of 0.15 M KCl. This fraction was applied to a 10 ml Heparin Sepharose CL-6B (Pharmacia) column, washed with 5 volumes of 0.35 M KCl and eluted with 2.5 volumes of 0.6 M KCl. This 0.6 M fraction was diluted to 0.1 M KCl, applied to a 5 ml DNA cellulose (native DNA; Pharmacia) column, and eluted with 2.5 volumes of 0.4 M KCl. This fraction was diluted to 0.15 M KCl and nonbinding mutant 16 (TGAGCCTTCTGC*GG*TGCGGGGACACCGTCT) was added at 10 ug/ml, incubated on ice 10 min and centrifuged at 12,000 x *g* for 10 min at 4°C to clear any precipitated material. The protein was divided in half and applied to two 1 ml DNA affinity columns. The columns were washed with 5 volumes of 0.5 M KCl and eluted with 5 volumes of 0.8 M KCl. The ERF-1 froaction was incubated with additional mutant oligonucleotide (10 *ug*/ml) and passed a second time over a single 1 ml DNA affinity column as describe above. All fractions were quick frozen in liquid nitrogen and

stroed at -80°C for subsequent analysis. The DNA affinity column was prepared by attaching a double-stranded, biotinylated 30-mer oligonucleotide (TGAGCCTTCTGCCCT GCGGGGACACGGTCT) corresponding to the wild-type ERF-1 binding site to streptavidin agarose.

Protein fractions were concentrated by trichloroacetic acid precipitation as described[22]. Protein pellets were resuspended in standard Laemmli SDS loading buffer, boiled 3 min, and electrophoresed on 8% SDS/PAGE with standard protein molecular weight markers. Proteins were visualized using the Silver Stain Plus kit (Bio-Rad).

Protein Sequencing

Protein was separated by SDS/PAGE, stained with Coomassie brilliant blue, and individual protein bands excised. After *in gel* tryptic (Promega) digestion as described [23] without the addition of 0.02% Tween. Candidate ions for sequencing were determined by microcapillary high performance liquid chromatography coupled to a triple mass spectrometer (model TSQ7000 with electrospray ionization source; Finnigan, San Jose, CA) as described [24]. The ion densities observed corresponded to a load of 40-100 fmol by comparison with average ion abundance of a BSA standard protein digest. Direct peptide sequence information was obtained by collisionally induced dissociation on an equivalent injection of the digest mixture. The resulting MS/MS spectra were manually interpreted. The data base searching algorithm SEQUEST [25] was also used to facilitate interpretation of MS/MS spectra.

UV crosslinking

DNA/protein crosslinking of the ERF-1 complex was performed as described [26] using 15 *u*g of MCF7 nuclear extract. Radiolabeled probe was prepared as described by annealing 1 pmol of an oligonucleotide encompassing the ERF-1 binding site

(CCGTGTCCCCGCAGGGCAGAAGGCTCA) (Operon Technologies) with 100 pmol of a complimentary oligonucleotide (TGAGCCTTCT) (Ana-Gen Technologies).

Renaturation of ERF-1 Activity from SDS/PAGE

Purified extract was resolved on 8% SDS/PAGE after which the gel was segmented into 4 mm slices ranging from 31 to 220 kDa. Each slice was crushed in 0.4 ml elution buffer (0.15 M NaCl/0.1% SDS/0.05 M Tris-HCl, pH 7.9/0.1 mM EDTA/5 mM DTT/0.1 mg/ml BSA) and protein was eluted for 1 h at room temperature, followed by precipitation with 4 volumes of cold acetone. The pellet was washed with cold 80% acetone. Precipitated protein was dissolved in 0.5 ml of 8 M urea in D-100 buffer and incubated at 4°C for 30 min. The protein was renatured by dialysis against 1 liter of 1 M urea in D-100 buffer for 3 h, followed by two changes of D-100 buffer for 3 h and 12 h, respectively.

Cloning of ERF-1 from an MCF7 cDNA Library

Primers AACCCTCTGAACCTCCCCTGTCAGAAG and CCGGTCTTGGCTGAG AAGTTCTGTGAATTC (Operon Technologies) were used to amplify a 504-bp fragment of the ERF-1 mRNA from MCF7 using the GeneAmp RNA PCR kit (Perkin-Elmer). This fragment was random primed labeled using [α-^{32}P]dCTP for use as a probe to retrieve a full-length ERF-1 cDNA clone from an MCF7 expression library prepared using the Zap Express cDNA Synthesis kit (Stratagene). Clones were excised from λ phage as phagemids and were analyzed by restriction analysis. ERF-1 cDNA was sequenced using an Applied Biosystems automated sequencer.

***In Vitro* Transcription and Translation**

ERF-1 was *in vitro* translated from ERF-1 cDNA using the TNT Coupled Reticulocyte Lysate System (Promega). Unlabeled ERF-1 was translated *in vitro* by using a complete amino acid mix in place of [^{35}S]methionine. For immunoprecipitations, 15 *u*l

of *in vitro* translated ERF-1 protein was incubated on ice for 10 min with 2 *u*g of affinity-purified rabbit polyclonal AP2 antibody in a 200 *u*l volume of RIPA buffer (150 mM NaCl/ 50 mM Tris, pH 8.0/ 1% Nonidet 40/ 0.5% deoxycholate/ 0.1% SDS). Fifty microliters of anti-rabbit IgG (whole molecule) agarose (Sigma) was added to the reaction and incubated on ice for 15 min. Immunoprecipitated protein was pelleted by centrifugation at 1,000 x *g* at 4°C for 5 min, washed once with RIPA buffer and twice each with buffer 1 (150 mM NaCl/ 10 mM Tris, pH 7.5/ 2 mM EDTA/ 1% Nonidet P-40) and buffer 2 (500 mM NaCl/ 10 mM Tris, pH 7.5/ 2 mM EDTA), and the final pellet was resuspended in standard Laemmli SDS loading buffer for SDS/PAGE analysis.

PCR-Assisted Binding-Site Selection

The technique of PCR-Assisted Binding-Site selection was used to determine the target specificity of AP2α and ERF-1. Random-sequence oligonucleotides to be used as a pool of potential binding sites were generated as described[27]. A 70 bp oligonucleotide containing a 20 bp random core (APbindSITE: CATTGATCTATTACAA GTCATACAGNNNNNNNNNNNNNNNNNNGAATCATGTTATTCAACAGTAA GAC) was annealed to a 25 bp oligonucleotide (APbindREV: GTCTTACTGTTGAATAA CATGATTC) complimentary to the 3' end. Oligonucleotides were labeled with [α-^{32}P] dCTP (800 Ci/mmol; Amersham) using Taq DNA polymerase (Perkin-Elmer) in a single cycle as follows: 1 min at 94°C; 3 min at 62°C; 9 min at 72°C. The extension reaction was chased with additional cold dCTP by incubating 9 min. at 72°C. The double-stranded probe was purified on 8% nondenaturing PAGE, excised from the gel and eluted in elution buffer (0.5 M ammonium acetate/ 1 mM EDTA/ 0.1% SDS) for 2 hours at 48°C. Eluate was removed to a fresh tube containing 20 ug glycogen and DNA was recovered by ethanol precipitation. The DNA pellet was resuspended in 10 ul dH_2O and 0.5 ul was counted in a scintillation counter to calculate recovery.

Twenty nanograms of GST protein, ERF-1-GST fusion protein, or affinity-purified AP2α (Promega) were incubated with 0.4 ng of labeled oligonucleotide pool and 1 μg AP2α antibody (184X; Santa Cruz Biochem, Santa Cruz, CA) in 1x binding buffer (40 mM KCl/ 20 mM Hepes, pH 7.9,/1 mM $MgCl_2$/ 0.08 mM EGTA/ 0.4 mM DTT/ 4% Ficoll/ 12 ug/ml poly dI-dC (Pharmacia)/ 1 mg/ml BSA (Sigma)). After the first round, binding reactions contained only 0.2 ng of labeled oligonucleotide pool. Reactions were incubated at RT for 30 min to allow protein-DNA complexes to form. Complexes were immunoprecipitated with protein A agarose (Gibco BRL) at 4°C overnight. Unbound oligonucleotides were removed by gently washing 3 times with 1x GSB (40 mM KCl/ 20 mM Hepes, pH 7.9,/1 mM $MgCl_2$/ 0.08 mM EGTA/ 0.4 mM DTT). Bound oligonucleotides were incubated in 200 ul recovery buffer (50 mM Tris·Cl, pH 8, 100 mM sodium acetate, 5 mM EDTA, 0.5% SDS) for 1 hr at 45°C, then extracted once each with phenol and chloroform. Specific binding site oligonucleotides were recovered by ethanol precipitation in the presence of 20 ug glycogen and 1/5 volume of 3 M sodium acetate. The pellet was resuspended in 5 ul of dH_2O and 0.5 ul was counted to calculate recovery.

One picogram of selected DNA was amplified by PCR in a reaction with 1X PCR buffer (50 mM potassium chloride/ 10 mM Tris-HCl, pH 8.3), 1 mM MgCl, 0.08 mM dGTP, 0.08 mM dATP, 0.08 mM dTTP, 4 μM dCTP, 0.01 mCi [α -^{32}P] dCTP (800 Ci/mmol), 160 ng each APbindFWD primer (CATTGATCTATTACAAGTCATACAG) and APbindRev primer, and 2.5 units Taq DNA polymerase (Perkin Elmer, Branchburg). Reactions were subjected to 18 PCR cycles as follows: 94°C, 1 min.; 50°C, 1 min.; 72°C, 1 min. DNA was gel purified and recovered as described above. Amplified DNA (0.2 ng) was used as input DNA for the next round of DNA-protein binding. This cycle of DNA binding and amplification was performed a total of 4 times.

To further refine the AP2 binding site, a second series of binding reactions were performed using a 57-mer oligonucleotide (CATTGATCTATTACAAGTCATACASCCNN

NGGSAATCATGTTATTCAACAGTAAGAC; S=G/C) as described, except only two binding/amplification cycles were performed.

Selected binding site oligonucleotides obtained after 4 binding/amplification cycles were ligated into pCR™II using the Original TA Cloning kit (Invitrogen Corp., San Diego, CA). Transformants were selected and plasmid DNA was prepared using the Wizard Plus Miniprep DNA Purification System (Promega, Madison, WI). Plasmid DNA was sequenced using the T7 Sequenase v2.0 DNA Sequencing Kit (Amersham).

To generate ERF-1 protein for PCR-assisted binding site selection, the ERF-1 cDNA was cloned into PGEX-4T-1 (Pharmacia Biotech, Piscataway, NJ) to create a fusion protein with glutathione S-transferase (GST). Using the ERF-1 clone retrieved from the MCF7 expression library as template, ERF-1 cDNA from the translation start site at +167 to the stop site at +1519 was PCR amplified using a 5' primer (GCCGCTCGAGATGTTG TGGAAAATAACCGA) and a 3' primer (TCGTTCTCGAGTTATTTCCTGTGTTCTCC), each of which contained an *Xho*I site. After digestion with *Xho*I, this fragment was ligated into the *Xho*I site of pGEX-4T-1 to generate ERF-1/pGEX-4T-1. ERF-1-GST fusion protein was induced and purified using Glutathione Sepharose 4B according to the protocol provided with the GST Gene Fusion System (Pharmacia Biotech). For use as a negative control, GST protein was induced from pGEX-4T-1 and purified in the same manner.

Competitive Gel Shift Assay

Competitive gel shift assays were performed to determine the relative binding affinity of a series of mutations in the optimal binding site for AP2α and ERF-1. Twenty-seven mutant double-stranded oligonucleotides were synthesized (Operon Technologies) with all possible point mutations in each position of the 9 bp optimal binding site, GCCTGAGGG. The mutant oligonucleotides were designed as 57-mers to contain the identical flanking sequence used throughout the PCR-Assisted Binding-Site selection (ie. Mut1A= CATTGATCTATTACAAGTCATACA**A**CCTGAGGGAATCATGTTATTCA

ACAGTAAGAC). The 57-mer containing the optimal binding site (OPT: GCCTGAGG G) was labeled as described for PCR-Assisted Binding-Site selection and 0.2 ng was used as probe in each reaction. Competitive gel shift assays were performed as described above except that protein amount was limited (0.2 ul AP2α; 0.5 ul ERF-1/GST fusion) to give a 50% decrease in binding with 25-75 fold excess of cold OPT to [α -^{32}P] dCTP-labeled OPT. For each point mutation, competitor was added in increasing amounts from 10- up to 5,000-fold excess over labeled probe.

Images were obtained by standard autoradiography and by use of a Molecular Dynamics (Sunnyvale, CA) phoshorimager. Quantitative values for binding were obtained from the phosphorimager using ImageQuant software (Molecular Dynamics). The fold excess of competitor at which a 50% decrease in binding occurred was determined using linear regression analysis. Values were corrected by dividing the value at which a 50% decrease in binding was obtained by the value at which a 50% decrease in binding occurred when OPT was competed with itself. This corrected value is referred to as Relative Binding Affinity.

Construction of AP2α/pcDNA3.1(+) and ERF-1/pcDNA3.1(+) Plasmids

To construct AP2α/pcDNA3.1(+), SP(RSV)AP2 (gift from Trevor Williams, New Haven, CT) was digested with *Kpn*I and *Xho*I to release a fragment containing the AP2α sequence from the translation start site at +62 through the 3' untranslated region. This fragment was subcloned into *Kpn*I/*Xho*I digested pcDNA3.1(+) mammalian expression vector (Invitrogen, San Diego, CA).

ERF-1/pcDNA3.1(+) was constructed by subcloning the *Xho*I fragment of ERF-1/pGEX-4T-1 into the *Xho*I site of pcDNA3.1(+). This *Xho*I fragment contains the ERF-1 cDNA from the start site at +167 to the stop site at +1519.

Construction of IL2-GL3/APWT and IL2-GL3/Mut3T Reporter Plasmids:

To determine the ability of AP2α and ERF-1 to activate an optimal binding site *versus* a mutated binding site *in vivo*, two luciferase reporter constructs were made. Two 59bp double-stranded, 5'-phosphorylated oligonucleotides with overhanging *Xho*I sites were synthesized (Operon Technologies) to contain either the optimal AP2 binding site (OPT: TCGAGATTGATCTATTACAAGTCATACAGCCTGAGGGAATCATGTTATTCAACA GTAAC) or a mutated binding site (M3T: TCGAGATTGATCTATTACAAGTCAT ACAGCTTGAGGGAATCATGTTATTCAACAGTAAC). These oligonucleotides were ligated into IL2-GL3 digested with *Xho*I to remove 90 bp containing 3 NFAT binding sites. IL2-GL3 (gift from Gerald Crabtree, Stanford, CA) is a luciferase reporter plasmid which contains the IL2 minimal promoter in addition to the 3 NFAT sites. DNA was prepared as described above and the insert was sequenced to verify sequence and direction of insertion.

Transfection

COS cells were plated at 720,000 cells/60 mm diameter culture dish the day before the assay to reach 80% confluency at the time of transfection. COS cells were transfected with a total of 6 µg DNA using LIPOFECTAMINE™ Reagent (GibcoBRL, Gaithersburg, MD). All transfections were performed in triplicate. Cells were harvested 24 hours post-transfection.

MDA-MB-231 and HepG2 cells were transfected using FuGENE™ 6 Transfection Reagent (Boehringer Mannheim,). Cells were plated the day before to reach 80% confluency at the time of transfection (MDA-MB-231: 7.2×10^5 cells/60 mm diameter culture dish; HepG2: 1.5×10^6 cells/60 mm diameter culture dish). Each plate was transfected with a total of 2.9 µg DNA and cells were harvested 24 hours post-transfection.

Whole cell extracts for gel shift assay or Western blot were prepared as described above. Cell extracts for luciferase assay or ß-Galactosidase assay were prepared using the

Luciferase Assay System With Reporter Lysis Buffer (Promega). Luciferase assays were performed with 20 µl of cell extract and 100µl of Luciferase Assay Buffer. Luciferase activity was measured for 10 sec in a Monolight® 2010 luminometer (Analytical Luminescence Laboratory, San Diego, CA). All plates were cotransfected with pßgal-Control Vector (Clontech, Palo Alto, CA) to measure transfection efficiency. ß-Galactosidase activity was assayed using the ß-Galactosidase Assay Kit (Stratagene, La Jolla, CA).

Western Blot

To demonstrate AP2α and ERF-1 expression in transfected cells, Western blot analysis was performed. Whole cell extract (7 µg/lane) was resolved on 8% SDS-PAGE, blotted to PVDF membrane and reacted with AP2α antibody (Santa Cruz Biotechnology) using the ECL Western Blotting Analysis System (Amersham Life Science Ltd, Buckinghamshire, England) as described.

Data Analysis

Luciferase data for AP2α and ERF-1 activation of IL2-pGL3, OPT/IL2 and M3T/IL2 were corrected for transfection efficiency by dividing the relative luciferase units by ß-Gal units for each transfection. Background activity of OPT/IL2 and M3T/IL2 was subtracted from the ß-Gal corrected luciferase units to yield Corrected Luciferase Units. Each transfection was performed in tripicate and data were analyzed for significance using ANOVA. Significance for two-group comparisons (OPT/IL2 *versus* IL2, M3T/IL2 *versus* IL2, and OPT/IL2 *versus* M3T/IL2) was tested using a Bonferroni *t* test [28]. Data are expressed as mean ± SD.

RESULTS

Functional Analysis of the ER Promoter

To elucidate the mechanisms which regulate transcription of the ER gene, we have performed a functional analysis of the ER promoter. A human genomic lambda library was screened with a probe from the 5' flanking region of the ER gene. A genomic clone was obtained which contained 3500 bp upstream of the P1 cap site for ER, the entire first coding exon, and approximately 10 kbp of the first intron. This 5' flanking region encompasses 500 bp upstream of the farthest ER cap site identified. Regions of this clone were then subcloned into the luciferase reporter vector pGL2-Basic. Figure 1 shows a diagram of the promoter constructs. The major ER mRNA beginning at P1 contains a 230 base untranslated 5' leader sequence [14]. In the first set of constructs, all inserts contained 210 bp of the untranslated leader and a nested set of 5' deletions were generated from 3500 bp down to the P1 cap site at +1 (Figure 1A.) In a second set of promoter constructs, the 5' end remained at 3500 bp and 3' deletions were constructed beginning at +230 and progressively deleting portions of the leader by sequentially bringing the luciferase gene closer to the P1 cap site (Figure 1B.)

These constructs were tested for luciferase expression upon transfection into two breast carcinoma cell lines (Figure 2). T47D is an ER-positive breast carcinoma line in which approximately 90% of the ER mRNA begins at the P1 cap site [15]. The ER-negative carcinoma line, MDA-MB-231 was also used since we have previously shown this cell line to lack transcription of the ER gene [11]. Luciferase activity in these two cell lines demonstrated strikingly different results (Figure 2). In ER-positive T47D, the full-length construct gave excellent expression. Progressive deletion of the 5' end of the gene failed to significantly alter expression, although there was a reproducible decline of expression upon deletion of the last 40 bp of the 5' flanking region which contains a TATA element at +25 (Figure 2A.) The pattern of expression in ER-negative MDA-MB-231 was qualitatively and quantitatively different. The full-length construct expressed poorly and gave values an

order of magnitude less than in T47D. This level of expression was only three times the expression from the negative control vector pGL2-Basic, which does not contain a promoter. Progressive deletion of the upstream sequences improved expression and no significant decline occurred when the TATA element upstream of the P1 cap site was deleted. Results from the 3' deletion constructs are shown in Figure 2B. Deletions from the 3' end of the leader identified a sequence between +210 and +135 which augmented expression of the ER promoter in ER-positive T47D cells. Analysis of these constructs in ER-negative MDA-MB-231 was also performed and the effect of this region appeared to be cell line specific. We, therefore, searched for binding proteins that interact with this region of the ER gene.

Identification of ERF-1 by Gel Shift

An 80 bp probe was prepared from sequences from +132 to +211 of the ER promoter. This probe was used in gel-shift assays with extracts prepared from the ER-positive breast carcinoma cell lines MCF7 and T47D and the ER-negative MDA-MB-231 cell line. A prominent shift band was found only in extracts from the ER-positive cell lines (Figure 3). This complex is referred to as ERF-1 (Estrogen Receptor Factor-1). These results indicate that a DNA binding protein, ERF-1 is expressed in ER-positive breast carcinoma cells and interacts with a region of the leader with transcriptional activity.

ERF-1 Expression in Breast and Endometrial Carcinoma

To determine which cell lines express ERF-1, a panel of human cell lines were analyzed for ERF-1 expression using the gel shift assay. These results are shown in Figures 4A and B. Abundant ERF-1 expression was found in all ER-positive breast carcinoma cell lines tested -MCF7, T47D, and BT20. Low levels of ERF-1 were detected in normal human mammary epithelial cells (HMEC). HBL-100 is an ER-negative breast carcinoma line which appears to express low amounts of ERF-1 comparable to HMEC.

The lack of expression of ER in HBL-100 could be due to any one of a number of reasons, e.g. methylation of DNA, deletion of the ER gene, rapid degradation of mRNA.

Expression of ERF-1 was also examined in other cell lines, including a panel of human endometrial carcinoma lines. The RL95-2 cell line is an endometrial adenocarcinoma line which is reportedly ER-positive [29]. This cell line makes abundant ERF-1 protein as seen in Figure 4B. Examination of ER expression in this stock of RL95-2 has failed to demonstrated ER expression (data not shown); however, late passage of this cell line has been reported to lose ER expression [30, 31]. Therefore, RL95-2 was derived from an ER-positive carcinoma and loss of ER expression as the cells are maintained in culture is likely the result of DNA methylation. ECC-1 is another endometrial cell line which is ER-positive and we have confirmed expression of ER mRNA (data not shown). The ECC-1 cell line also expresses abundant ERF-1 (Figure 4B). ERF-1 was not readily detected in HEC 1B or HEC 1A, which are both ER-negative human endometrial carcinoma cell lines [32]. These results suggest that abundant expression of ERF-1 represents a common mechanism for ER regulation in hormonally responsive carcinomas.

Mapping the ERF-1 Binding Site.

Gel shift competitions were used to identify two ERF-1 binding sites within the 80 bp region (+132 to +211) of the ER promoter shown to bind ERF-1. Cold competitor prepared from the entire 80 bp probe efficiently competes for binding (Figure 3). The region from +132 to +171 partially competes and +172 to +211 competes efficiently. Within the region from +172 to +211, the sequences from +182 to +201 compete efficiently for binding (Figure 3). Neither +172 to +191 nor +192 to +211 demonstrate any competition. Therefore, ERF-1 binds to two sites in this region, a distal (high affinity) and a proximal (low affinity) site.

Mutations of ERF-1 Binding Sites affects Binding

The distal binding site was mapped more precisely utilizing gel shift competition with oligonucleotides containing mutations within the sequences from +182 to +201. These results are shown in Figure 5A. This region contains the sequence CCCTGCGGGG which is an imperfect palindrome. The wild-type sequence of this distal site (dwt) competes efficiently. Mutations d1 and d3 disrupt the imperfect palindrome and destroy the ability of the oligonucleotide to compete. Mutations d2 and d4 do not alter the 10 bp imperfect palindrome and retain the ability to compete for binding although d4 is slightly less efficient than dwt. Mutation d5 changes the T at +192 to a C and converts the sequence to a perfect palindrome but partially weakens the ability to efficiently compete. The sequence between +132 to +171 that demonstrated weak competition shown in Figure 3, was found to contain a second ERF-1 site located at +130 to +149 (Figure 5B). Within this region, a related sequence can be found and homologous mutations as used for the distal site have identical consequences for binding. Interestingly, insertion of a G between +140 and +141 creates a site identical to the distal site. This mutation improves the ability of the weak proximal site to compete.

Mutation of ERF-1 Sites Affects Expression

Because the mutation p1 destroys binding to the proximal site and d1 destroys binding to the distal site, these two mutations were built into the expression vector ER3500-230LUC and called ER3500-230p1d1. This new vector is identical to ER3500-230LUC except for the two mutations within the proximal and distal ERF-1 binding sites. Expression from this construct is shown in Figure 2B. Mutation of these ERF-1 sites has an effect on expression similar to deletion of the region from +135 to +210. These results strongly suggest that ERF-1 is a transcription factor which is expressed in ER-positive breast carcinoma and which functions by binding to two sites in the untranslated leader of the ER gene.

Mapping of the ERF-1 Distal Binding Site

More precise mapping of the distal site was performed by gel shift competition using oligonucleotides with double point mutations within the sequences +178 to +207. This sequence contains the imperfect palindrome, CCCTGCGGGG. Results are seen in Figure 6. Mutants 7,8,9, and 10 disrupt the imperfect palindrome and fail to compete for binding to the wild-type probe. Mutations that target several nucleotides flanking the palindrome demonstrate partial binding specificity. The 30 bp mutants 1,2,13,14, and 15 demonstrate competition equal to wild-type probe. These results confirm that the ERF-1 binding site is represented by the imperfect palindrome with some sequence specificity demonstrated by flanking nucleotides.

Purification of ERF-1.

ERF-1 protein was purified from MCF7 nuclear extract using ion exchange and DNA affinity chromatography. The 30 bp double stranded DNA probe comprising sequences from +178 to +207 was used for the DNA affinity step. A diagram of the purification protocol is shown in Figure 7. ERF-1 activity was followed through the purification process using gelshift analysis.

Molecular Weight Determination of ERF-1

Both UV crosslinking and gel renaturation experiments indicated the size of ERF-1 as ≈ 50 kDa. First, the bound ERF-1 complex resolved by gel shift was UV crosslinked to a radiolabeled ERF-1 DNA probe. This protein-DNA complex had an apparent molecular weight of 60-65 kDa on SDS/PAGE as shown in Figure 8A. Subtracting the molecular weight of the DNA would indicate a protein of ≈ 50-55 kDa. Second, ERF-1 activity was renatured from SDS/PAGE. Purified extract was electrophoresed on 8% SDS/PAGE as shown in Figure 8B. This purified extract contained two major bands, as well as several minor species. As can be seen in Figure 8C, fraction number 7 contains ERF-1 activity.

This renatured complex comigrates with native ERF-1 complex It has also been shown that this renatured complex demonstrates identical binding specificity as determined for the native protein as shown in Figure 6 (data not shown). All other fractions are negative for binding activity. Fraction number 7 corresponds to a region of the gel of ≈ 50 kDa, which agrees with the size determined by crosslinking studies.

Fraction number 7 of the purified extract contained two obvious protein bands. Each of these bands was excised and subjected to peptide sequencing. The lower, more prominent band was identified as La/SS-B, which is an RNA binding protein [33]. La/SS-B was unlikely to be ERF-1 based upon known characteristics of this protein. In addition, three separate antibodies to La/SS-B failed to supershift native ERF-1 complex (data not shown). Sequence analysis of the upper, less prominent band yielded two peptide sequences, NPLNLPCQK and EFTELLSQDR. Both of these peptides matched precisely the amino acid sequence of a partial cDNA clone for AP2γ (GenBank accession no. X95693). AP2γ was identified based upon homology to the AP2α transcription factor [34].

Isolation of ERF-1 cDNA

Two DNA oligonucleotide primers were prepared corresponding to the two peptide sequences obtained from peptide sequence analysis. These primers were used to amplify MCF7 mRNA utilizing RT-PCR, which generated an expected PCR product of 504 bp (data not shown). The PCR product was used as a probe to screen an MCF7 cDNA library. Thirteen separate cDNA clones were obtained, with the nine largest clones having an identical pattern on restriction analysis. The size of these cDNA inserts was 2.7 kb. One of these cDNAs was sequenced completely and this sequence is shown in Figure 9A. This cDNA contains a long open reading frame of 1353 bp encoding a protein with a predicted molecular weight of 48 kDa. The predicted amino acid sequence of the ERF-1 protein matched precisely the amino acid sequence reported for AP2γ. The ERF-1 cDNA contains a 166-bp 5' untranslated leader and a 1268-bp 3' untranslated region. Alignment of the

ERF-1 protein with AP2α is shown in Figure 9B. Overall, there is a 65% identity and 83% similarity between these two proteins. Homology is most striking within the carboxyl-terminal half of the proteins which contains the DNA binding and dimerization domains. Within the carboxyl-terminal half there is 76% identity and 90% similarity between these two proteins.

Confirmation of ERF-1 Identity. The identity of ERF-1 was confirmed by analysis of *in vitro* synthesized product and by antibody reactivity. The ERF-1 cDNA was subjected to *in vitro* transcription/translation. As can be seen in Figure 10A, the *in vitro* product has a molecular weight of ≈50 kDa. A polyclonal antisera generated against the carboxyl-terminal 20 amino acids of AP2α is able to immunoprecipitate the *in vitro* synthesized product. The T7 promoter is directed in the opposite orientation and, as expected, no product is obtained. *In vitro* synthesized ERF-1 was prepared in the absence of ^{35}S-labeled amino acids. This cold protein was analyzed in a gel shift assay as shown in Figure 10B. *In vitro* synthesized ERF-1 generates an ERF-1 complex that comigrates with the native ERF-1 complex. This product also has the same DNA sequence specificity as demonstrated with competition using mutant oligonucleotide sequences. In addition, AP2α antisera supershifts the *in vitro* ERF-1 complex. As expected, the extract prepared using the T7 polymerase does not generate the ERF-1 complex. The AP2α antisera should also identify the native ERF-1 complex. As can be seen in Figure 11, this antisera supershifts all of the ERF-1 complex from MCF7 cells This supershifted complex also demonstrates the same pattern of sequence binding as determined by mutant competitors.

It is evident that the ERF-1 binding sequence is similar to the sequence specificity of AP2α (GCCNNNGGC) [35] and AP2α antisera cross reacts with ERF-1 protein. Given the extensive homology between AP2 family members, it might be questioned whether the ERF-1 complex is formed by AP2α homodimers. However, several considerations make this possibility unlikely. First, HeLa cells express AP2α whereas ERF-1 activity was not

readily demonstrated in this cell line and several breast carcinoma cell lines known to express ERF-1 do not express AP2α [36, 37]. Second, although cloned AP2α protein binds to the wild-type ERF-1 probe by gelshift analysis, the complex formed demonstrates different binding specificities for mutant competitors and has slightly faster mobility as compared with native ERF-1 complex. Finally, the peptides obtained from sequencing purified ERF-1 differ from the amino acid sequence for AP2α. These findings indicate that ERF-1, previously shown to be involved in regulation of the ER gene promoter, is an AP2 family member distinct from AP2α.

PCR-Assisted Binding Site selection

PCR-Assisted binding site selection identified the preferred binding site for AP2α as $GCC^{C}/G/_{T}{}^{G}/C/_{T}{}^{G}/A/_{C}{}^{G}/_{A}G^{G}/_{C}$ (Table 1) and that for ERF-1 as $GCC^{C}/T/_{G}{}^{G}/_{C}{}^{G}/A/_{C}GG^{G}/_{C}$ (Table 2). To provide a starting point, binding site clones were searched for sequences beginning with GCC as described previously for the AP2α consensus site (GCCNNNGGC) [35]. Of 132 selected clones sequenced for AP2α, 108 clones contained one or more binding sites beginning with GCC, 5 clones contained no binding sequence fitting this criteria, and 19 clones contained unreadable sequence. For 129 clones sequenced for ERF-1, 92 clones contained one or more binding sites, 8 clones contained no binding site, and 29 clones had unreadable sequence.

Many clones contained multiple binding sites, either on the same and/or on the complementary strand, making it impossible to know which sequence was primarily responsible for selection of the clone. Because this may have biased the results, the sequences were also analyzed to identify clones in which only a single binding site appeared. This analysis created some changes in preferred binding site sequence, most notably at positions 5 and 6 for AP2α and at position 7 for ERF-1 (Tables 3 and 4).

To further refine the binding site for AP2α and ERF-1, PCR-Assisted Binding Site selection was repeated using a probe containing the binding site SCCNNNGGS (S=G/C).

This selection cycle defined the preferred AP2α binding site to be $GCC^{C}/G/_{T}GNGG^{G}/_{C}$ (Table 5) and the preferred ERF-1 binding site to be $GCC^{T}/C/_{G}G^{A}/G/_{C}GG^{G}/_{C}$ (Table 6). Of 36 clones sequenced for each protein, readable sequences were obtained for 27 AP2α clones and for 25 ERF-1 clones.

Competitive Gel Shift Assay

Competitive gel shift assays using a series of competitors with all possible single point mutations in the optimal binding site demonstrated that there were no significant differences in binding site preference between AP2α and ERF-1 (Figures 12 and 13). Assuming that any competitor with a Relative Binding Affinity of 6 or lower bound protein with equal affinity as OPT (GCCTGAGGC), the preferred binding site for both AP2α and ERF-1 was determined to be $^{G}/_{C}CCNN^{A}/C/_{G}{}^{G}/_{A}{}^{G}/_{A}{}^{G}/C/_{T}$.

Transfections of COS cells

Cotransfection of COS cells with AP2α or ERF-1 and luciferase reporter constructs containing either the optimal binding sequence (OPT/IL-2) or a mutated binding sequence (M3T/IL-2) demonstrated that both proteins activated OPT/IL-2 to a significantly higher level than M3T/IL-2 over a range of protein concentrations (Figures 14C and 15C). For AP2α, activation of OPT/IL-2 was significantly higher than M3T/IL-2 when transfections were performed with 0.1, 0.5, 1, or 2 μg protein. There was no significant differences in activation at 0, 0.2, or 3 μg. For ERF-1, activation of OPT/IL-2 was significantly higher than M3T/IL-2 at all protein concentrations at or above 0.2 μg. For AP2α, the levels of activation of both OPT/IL-2 and M3T/IL-2 were significantly higher than activation of a reporter construct containing only the IL-2 minimal promoter (IL-2) except at 1 ug for M3T/IL-2 *vs* IL-2. For ERF-1, activation of both OPT/IL-2 and M3T/IL-2 were significantly higher than IL-2 at all protein concentrations above 0 μg. Overall, activation of both OPT/IL-2 and M3T/IL-2 by ERF-1 was 3.3 fold higher than activation by AP2α.

Western blot analysis and gel shift assay show that expression and DNA binding of AP2α and ERF-1 increases linearly as COS cells are transfected with increasing amounts of these proteins (Figure 14A and B; Figure 15A an B).

Transfection of HepG2 and MDA-231 Cells

Because ERF-1 was a better activator of OPT/IL-2 and M3T/IL-2 in COS cells than AP2α, activation of a full-length ER promoter construct (ER3500-230LUC) by ERF-1 and AP2α was tested in HepG2 and MDA-231, two cell lines lacking expression of these proteins. As described above, the ER promoter contains a 10 bp imperfect palindrome (CCCTGCGGGG) that acts as a binding site for ERF-1. To fit into the 9 bp pattern determined for optimal DNA binding, this sequence can be viewed as either CCCTGCGGG or CCTGCGGG, the latter being equivalent to the Mut 3T mutation. Preliminary data showed that in HepG2 cells, the activation of ER3500-230LUC by both AP2α and ERF-1 is greater than activation of PGL2-Basic which contains no promoter (Figure 16A). However, in MDA-MB-231 cells, only ERF-1 is able to activate ER3500-230LUC above background levels (Figure 16B).

DISCUSSION

Although several mechanisms are involved in controlling expression of ER in breast carcinomas, transcriptional regulation is clearly responsible for the ER-negative phenotype in some cell lines [11, 38]. These data demonstrate that an important transcriptional regulatory element of the ER gene is found within the 5' untranslated leader. This region contains two binding sites for a DNA binding protein, ERF-1, which is abundantly expressed in ER-positive breast and endometrial carcinomas. These data indicate a role for ERF-1 in the transcriptional regulation of ER. There are likely other cis-acting control elements involved in the regulation of ER transcription. For example, the transcriptional

mapping data also implicate positive and negative regulatory elements upstream of the P1 transcriptional start site and there may be other regions that have not been examined in the present study. However, the identification of ERF-1 offers a molecular mechanism that accounts for differences in ER expression found in breast carcinomas.

The complexity of ER transcriptional regulation is suggested by the biology of ER expression. ER is over-expressed in many ER-positive breast carcinomas that often make over 100 fmol/mg cytosol protein compared to 4 fmol/mg cytosol protein in normal mammary cells [2]. These differences in ER expression are reflective of mRNA levels and there are striking differences in the level of ER mRNA detected in different breast carcinoma lines [15]. It is certainly possible that ERF-1 may be responsible for ER over-expression identified in many ER-positive carcinomas. However, more detailed experiments need to be performed to address this possibility directly.

During this study, we have now cloned the gene for ERF-1 and have established that ERF-1 is a member of the AP2 transcription factor family. Three members of this family have been identified, AP2α [35], AP2ß and AP2γ [34]. The protein sequence for ERF-1 matches precisely the predicted amino acid sequence for AP2γ. The identification of ERF-1 as a member of a well-characterized transcription factor family provides additional proof supporting the role of ERF-1 in ER gene regulation. However, it remains to be determined if ERF-1 is responsible for basal tissue-specific expression of ER or if ERF-1 induces the over-expression of ER that is characteristic of many ER-positive carcinomas. Identification of the ERF-1 gene now facilitates the resolution of these mechanistic issues.

The homology of ERF-1 with AP2α raises several points of speculation concerning the function and biology of ERF-1. AP2α activates transcription by binding as a homodimer to specific DNA sequences [35] It has also been shown that *in vitro* synthesized AP2α, AP2γ and AP2ß can form heterodimers with one another and can bind to an AP2α consensus site [39]. It is not known if heterodimers between AP2 family members occur *in vivo* or whether they are physiologically relevant. However, heterodimer formation

between ERF-1 and other members of the AP2 family raises an additional potential for complex gene regulation. The formation of heterodimers may alter normal function of AP2 proteins or, alternatively, may generate complexes with different specificity. Germane to gene regulation in breast carcinomas, recent reports have implicated AP2 factors in the transcriptional control of c-*erb*B-2 [37] E-cadherin [40] and HSP27 [41]. It is certainly plausible to hypothesize a role for ERF-1 in the coordinate regulation of a set of genes in hormonally responsive carcinomas. This conjecture is supported by the fact that breast carcinoma cell lines that express E-cadherin and HSP27 are also ERF-1 positive, whereas cell lines with low level expression of these genes are ERF-1 negative [36, 40, 41]. Identifying other genes regulated by ERF-1 will be an important area for futher study.

Using PCR-Assisted Binding Site selection in conjunction with competitive gel shift assay, this study also demonstrates that AP2α and ERF-1 have identical DNA binding specificities. Previously, as determined by methylation interference assay and missing contact probing assay, the consensus binding site for AP2α was determined to be $GCCN_3GGC$ [35]. An earlier study had derived an slightly different AP2α consensus site using DNA footprinting studies of four genes (hGH, H-$2K^b$, hc-*myc*, and BPV) to have a T in the first position of a 10 bp site ($^{T}/_{C}C^{C}/_{G}CC^{A}/_{C}N^{G}/_{C}C/_{G}G/_{C}$) [42]. However, this consensus site was derived from a very small sample of naturally occurring AP2α binding sites and no experiments were performed to determine the affinity of the site or the relative importance of each nucleotide. Although it has not been formally tested, it was assumed that AP2α, AP2γ and AP2ß all had the same binding specificity due to the fact that they can bind to a high affinity AP2α consensus site in a gel shift assay and are able to activate a reporter construct containing three tandem copies of this site [39]. However, in the same study it was also demonstrated that AP2ß activated a reporter construct containing the AP2 binding site from the c-*erb*B-2 3-fold less than did AP2α and AP2γ, indicating that there may be some subtle differences in binding specificity between these proteins. To our knowledge, a systematic study in which each position of the DNA binding site was

thoroughly mutated had not been previously undertaken. Although for the most part this study confirmed that the optimal binding site for AP2α and ERF-1 fits within the previously described consensus site, we were surprised to find that a mutated binding site with a T in the position 9 had the same relative binding affinity as a site containing either a G or C. In addition, due to the palindromic nature of the consensus site, it is surprising that a mutated binding site with an A in position 1 destroyed its ability to bind both AP2α or ERF-1.

If there are no differences in binding specificity for AP2α and ERF-1, then any potential functional differences exhibited by these transcription factors must be due to other factors such as spatially or temporally restricted expression patterns or differences in the cohort of cofactors required for activation. In support of this, it was shown in mouse embryos that there are distinct patterns of AP2α and AP2ß expression in the developing brain and face [43]. Additionally, Western and Northern blot analysis of a panel of human mammary carcinoma cell lines also indicated overlapping but distinct expression patterns for AP2α, AP2ß, and AP2γ, with a 4-5-fold amplification of the AP2γ gene in two cell lines, BT474 and SKBR3 [39]. Several studies suggest that competition from adjacent DNA-binding proteins is able to displace AP2α from its binding site [44-46]. However the analagous experiments have not been done for AP2γ and to date, no interacting cofactors for either AP2α or AP2γ have been identified.

The validity of selecting an optimal binding site using PCR selection and competitive gel shift assay was examined *in vivo* by testing the activation of reporter constructs containing either the optimal binding site or a mutated site with a T in position 3 (M3T). This mutation was determined to be of lower affinity by competitive gel shift assay. This experiment shows that a reporter construct containing the M3T mutated site is activated 2.2-fold less than the construct containing the optimal binding site at certain amounts of imput AP2α (Figure 14C) and 1.4-fold less with ERF-1 (Figure 15C). For AP2α, the difference in activation of the optimal and mutated sites is not significantly

different with 3 ug of imput AP2α. It is possible that when AP2α is overexpressed to certain levels, the distinction between low and high affinity binding sites is negated. AP2α and ERF-1 expression in normal breast epithelium is very low [36] while in certain mammary carcinoma cell lines the expression of one or both of these proteins is greatly overexpressed [36, 39]. Therefore, it is possible that overexpression of AP2α or ERF-1 in breast cancer may activate a set of genes containing lower affinity binding sites that would normally not be activated in normal breast epithileum.

Preliminary experiments in HepG2 and MDA-MB-231 cells, two cell lines lacking AP2 activity, show that the ability of AP2α or ERF-1 to activate a reporter construct containing the ER promoter is dependent on cell type (Figure 16A and B). In HepG2 cells, a hepatic cell line, both AP2α and ERF-1 are able to activate the ER promoter construct above background levels with AP2α demonstrating a 2.5-fold greater activation than ERF-1. However, in MDA-MB-231 cells, an ER-negative breast carcinoma cell line, only ERF-1 is able to activate the ER promoter construct. It is possible that there are breast cell-specific factors present in MDA-MB-231 cells that are absent in a hepatic cell such as HepG2 which repress AP2α activation while allowing ERF-1 activation of the ER promoter construct to occur.

CONCLUSIONS

The identification and cloning of ERF-1 offers new insight into our understanding of the relationship between ER expression and the biology of breast and endometrial carcinoma. ER expression defines a subset of breast cancer patients who, in general, have a better prognosis compared to patients with ER-negative tumors. Because ER is a transcription factor, it has been suggested that the phenotype displayed by ER-positive breast carcinomas is due to the repertoire of genes whose expression is regulated through estrogen response elements (ERE). Alternatively, ER expression might be a marker for the degree of differentiation of a tumor and ERF-1 might be involved in the regulation of a

number of cellular genes, including ER, which are critical to the differentiated phenotype. The identification of ERF-1 has immediate clinical relevance. For example, tumors that lack ERF-1 expression might define a subset of cancer patients with a prognosis different than patients with ER-negative tumors where loss of expression is due to mutations within the ER gene. If ERF-1 expression is a more reliable marker of a clinically relevant phenotype, this would indicate that some ERF-1 responsive gene, other than ER, is critical to the phenotype of ER-positive carcinomas. Understanding the control of ERF-1 may also provide new therapeutic approaches to the treatment of aggressive ER-negative tumors.

REFERENCES

1. Iglehart, J.D. in *Textbook of Surgery: The biological basis of modern surgical practice* (eds. Sabiston, D.C.) 510 (W. B. Saunders Company, Philadelphia, 1986).
2. Ricketts, D., *et al.* Estrogen and progesterone receptors in the normal female breast. *Can. Res.* **51**, 1817-1822 (1991).
3. Knight, W.A., Livingston, R.B., Gregory, E.J. & McGuire, W.L. Estrogen receptor as an independent prognostic factor for early recurrence in breast cancer. *Cancer Res.* **37**, 4669-4671 (1977).
4. Clark, G.M., *et al.* How to integrate steroid hormone receptor, flow cytometric, and other prognostic information in regard to primary breast cancer. *Cancer sup.* **71**, 2157-2162 (1993).
5. Watts, C.K.W., Handel, M.L., King, R.J.B. & Sutherland, R.L. Oestrogen receptor gene structure and function in breast cancer. *J. Steroid Biochem. Molec. Biol.* **41**, 529-536 (1992).
6. Scott, G.K., Kushner, P., Vigne, J.-L. & Benz, C.C. Truncated forms of DNA-binding estrogen receptors in human breast cancer. *J. Clin. Invest.* **88**, 700-706 (1991).
7. Ince, B.A., Zhuang, Y., Wrenn, C.K., Shapiro, D.J. & Katzenellenbogen, B.S. Powerful dominant negative mutants of the human estrogen receptor. *J. Bio. Chem.* **268**, 14026-14032 (1993).
8. Fuqua, S.A.W., *et al.* Inhibition of estrogen receptor action by a naturally occurring variant in human breast tumors. *Can. Res.* **52**, 483-486 (1992).
9. McGuire, W.L., Chamness, G.C. & Fuqua, S.A.W. Estrogen receptor variants in clinical breast cancer. *Mol. Endocr.* **5**, 1571-1577 (1991).

10. Castles, C.G., Fuqua, S.A.W., Klotz, D. & Hill, S.M. Expression of a constitutively active estrogen receptor variant in the estrogen receptor-negative BT-20 human breast cancer cell line. *Can. Res.* **53**, 5934-5939 (1993).

11. Weigel, R.J. & deConink, E.C. Transcriptional control of estrogen receptor in estrogen receptor-negative breast carcinoma. *Cancer Res.* **53**, 3472-3474 (1993).

12. Keaveney, M., *et al.* Evidence for a previously unidentified upstream exon in the human oestrogen receptor gene. *J. Mol. Endo.* **6**, 111-115 (1985).

13. Thompson, D.A., McPherson, L.A., Carmeci, C., deConinck, E.C. & Weigel, R.J. Identification of two estrogen receptor transcripts with novel 5' exons isolated from a MCF7 cDNA library. *J. Steroid Biochem. Molec. Biol.* **62**, 143-153 (1997).

14. Green, S., *et al.* Human oestrogen receptor cDNA: sequence, expression and homology to V-*erb*-A. *Nature* **320**, 134-139 (1986).

15. Weigel, R.J., Crooks, D.L., Iglehart, J.D. & deConinck, E.C. Quantitative analysis of the transcriptional start sites of estrogen receptor in breast carcinoma. *Cell Growth and Differentiation* **6**, 707-711 (1995).

16. White, R., Lees, J.A., Needham, M., Ham, J. & Parker, M. Structural organization and expression of the mouse estrogen receptor. *Mol. Endo.* **87**, 735-744 (1987).

17. Piva, R., Bianchi, N., Aguiari, G.L., Gambar, R. & Del Senno, L. Sequencing of an RNA transcript of the human estrogen receptor gene: evidence for a new transcriptional event. *J. Steroid Biochem. Molec. Biol.* **46**, 531-538 (1993).

18. Band, V. & Sager, S. Distinctive traits of normal and tumor-derived human mammary epithelial cells expressed in a medium that supports long-term growth of both cell types. *Proc. Natl. Acad. Sci. USA* **86**, 1249-1253 (1989).

19. Sambrook, J., Fritsch, E.F. & Maniatis, T. *Molecular cloning: a laboratory manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989).

20. Graham, F.L. & van de Eb, A.J. A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* **52**, 456-457 (1973).

21. Dignam, J.D., Lebovitz, R.M. & Roeder, R.G. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* **11**, 1475-1489 (1983).

22. Peterson, G.L. A Simplification of the Protein Assay Method of Lowry et al. Which is More Generally Applicable. *Anal. Biochem.* **83**, 346-356 (1977).

23. Hellman, U., Wernstedt, C., Gonez, J. & Heldin, C.-H. Improvement of an "In-Gel" digestion procedure for the micropreparation of internal protein fragments for amino acid sequencing. *Anal. Biochem.* **224**, 451-455 (1995).

24. Hunt, D.F., *et al.* Characterization of peptides bound to the class I MHC molecule HLA-A2.1 by mass spectrometry. *Science* **255**, 1261-1263 (1992).

25. Eng, J.K., McCormick, A.L. & Yates, R., III. *J. Am. Soc. Mass Spectrom.* **5**, 976-989 (1994).

26. Ballard, D.W., *et al.* Activation of the interleukin-2 receptor alpha gene: regulatory role for DNA-protein interactions flanking the kappa B enhancer. *New Biol.* **1**, 83-92 (1989).

27. Pollock, R.M. in *Current Protocols in Molecular Biology* (eds. Kingston, R.E.) 12.11.1-12.11.8 (John Wiley & Sons, Inc, Boston, MA, 1996).

28. Stanton, A.G. *Primer of Biostatistics* 1-473 (McGRAW-HILL, New York, 1997).

29. Way, D.L., Grosso, D.S., Davis, J.R., Surwit, E.A. & Christian, C.D. Characterzation of a new human endometrial carcinoma (RL95-2). *In Vitro* **19**, 147-158 (1983).

30. Sundareshan, P. & Hendrix, J.C. Growth, morphologic, and invasive characteristics of early and late passages of a human endometrial carcinoma cell line (RL95-2). *In Vitro Cell Dev. Biol.* **28A**, 544-552 (1992).

31. Kleinman, D., Douvdevani, A., Schally, A.V., Levy, J. & Sharoni, Y. Direct growth inhibition of human endometrial cancer cells by the gonadotropin-releasing

hormone antagonist SB-75: role of apoptosis. *Am. J. Obstet. Gynecol.* **170**, 96-102 (1994).

32. Kuramoto, H., Tamura, S. & Notake, Y. Establishment of a cell line of human endometrial adenocarcinoma in vitro. *Am. J. Obstet. Gynecol.* **114**, 1012-1019 (1972).
33. Bachman, M., Pfeifer, K., Schröder, H.C. & Müller, W.E.G. Characterization of the Autoantigen La as a Nucleic Acid-Dependent ATPase/dATPase with Melting Properties. *Cell* **60**, 85-93 (1990).
34. Williamson, J.A., *et al.* Chromosomal Mapping of the Human and Mouse Homologues of Two New Members of the AP-2 Family of Transcription Factors. *Genomics* **35**, 262-264 (1996).
35. Williams, T. & Tijian, R. Analysis of the DNA-binding and activation properties of the human transcription factor AP-2. *Genes Dev.* **5**, 670-682 (1991).
36. de Coninck, E.C., McPherson, L.A. & Weigel, R.J. Transcriptional Regulation of Estrogen Receptor in Breast Carcinoma. *Mol. Cell. Biol.* **15**, 2191-2196 (1995).
37. Bosher, J.M., Williams, T. & Hurst, H.C. The developmentally regulated transcription factor AP-2 is involved in *c-erfB-2* overexpression in human mammary carcinoma. *Proc. Natl. Acad. Sci. USA* **92**, 744-747 (1995).
38. Ottaviano, Y.L., *et al.* Methylation of the estrogen receptor gene CpG island marks loss of estrogen receptor expression in human breast cancer cells. *Can. Res.* **54**, 2552-2555 (1994).
39. Bosher, J.M., Totty, N.F., Hsuan, J.J., Williams, T. & Hurst, H.C. A family of AP-2 proteins regulates c-*erb*B-2 expression in mammary carcinoma. *Oncogene* **13**, 1701-1707 (1996).
40. Hennig, G., Löwrick, O., Birchmeier, W. & Behrens, J. Mechanisms Identified in the Transcriptional Control of Epithelial Gene Expression. *J. Biol. C hem.* **271**, 595-602 (1996).

41. Oesterreich, S., Hickey, E., Weber, L.A. & Fuqua, S.A.W. Basal regulatory promoter elements of the hsp27 gene in human breast cancer cells. *Biochem. Biophys. Res. Com.* **222**, 155-163 (1996).

42. Imagawa, M., Chiu, R. & Karin, M. Transcription Factor AP-2 mediates Induction by Two Different Signal-Transduction Pathways: Protein Kinase C and cAMP. *Cell* **51**, 251-260 (1987).

43. Moser, M., *et al.* Cloning and Characterization of a second AP-2 transcription factor: AP-2ß. *Development (Cambridge, UK)* **121**, 2779-2788 (1995).

44. Israël, A., *et al.* TNF stimulates expression of mouse MHC class I genes by inducing an NFkB-like enhancer binding activity which displaces constitutive factors. *EMBO j.* **8**, 3793-3800 (1989).

45. Mercurio, F. & Karin, M. Transcription factor AP-3 and AP-2 interact with the SV40 enhancer in a mutually exclusive manner. *EMBO j.* **8**, 1455-1460 (1989).

46. Courtois, S.J., Lafontaine, A., Lemaigre, F.P., Durviaux, S.M. & Rousseau, G.G. Nuclear factor-1 and activating protein-2 bind in a mutually exclusive way to overlapping promoter sequences and trans-activate the human growth hormone gene. *Nucleic Acids Res.* **18**, 57-64 (1990).

FIGURE LEGENDS

Figure 1. ER promoter constructs.

Diagram of ER promoter constructs cloned in luciferase expression vector pGL2-Basic. Dark line represents DNA sequence derived from ER gene. Arrow shows location of P1 transcriptional start site. All numbers correspond to distance from P1 cap site. Location of *Nde*I site at -45 and *Xho*I site (in vector) are shown. ER coding region is indicated in box. Luciferase coding region is designated by box labelled LUC. In plasmid names, first number corresponds to DNA length upstream of P1 and second number is DNA length downstream of P1. A. 5' deletion constructs are shown. All these constructs contain 210 bp of the 230 bp 5' untranslated leader. B. 3' deletion constructs are shown. All the constructs contain 3500 bp upstream of P1 and variable portions of the untranslated leader.

Figure 1A.

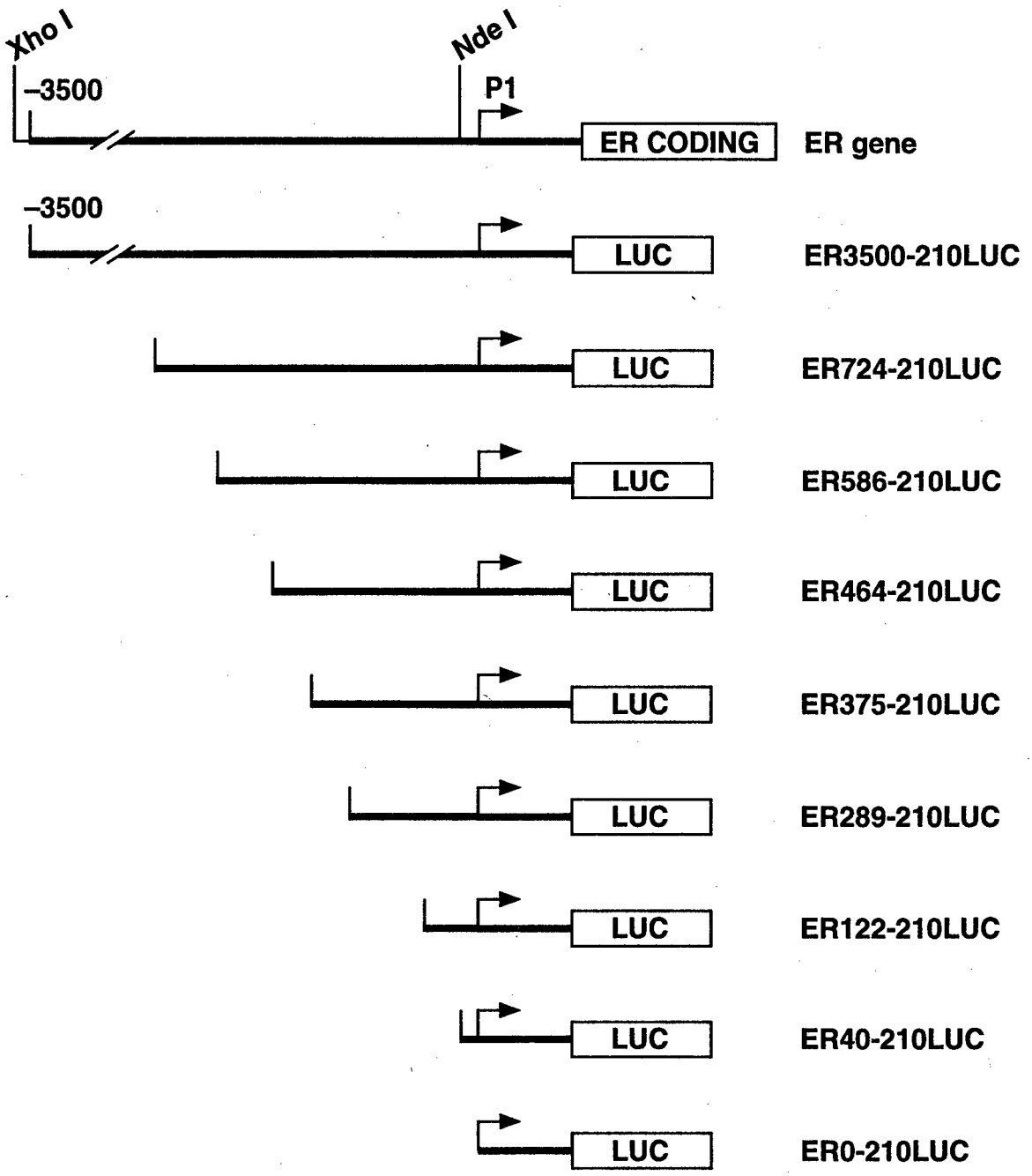

Xho I
Nde I
-3500
P1
ER CODING
ER gene
-3500
LUC
ER3500-210LUC
LUC
ER724-210LUC
LUC
ER586-210LUC
LUC
ER464-210LUC
LUC
ER375-210LUC
LUC
ER289-210LUC
LUC
ER122-210LUC
LUC
ER40-210LUC
LUC
ER0-210LUC

Figure 1B.

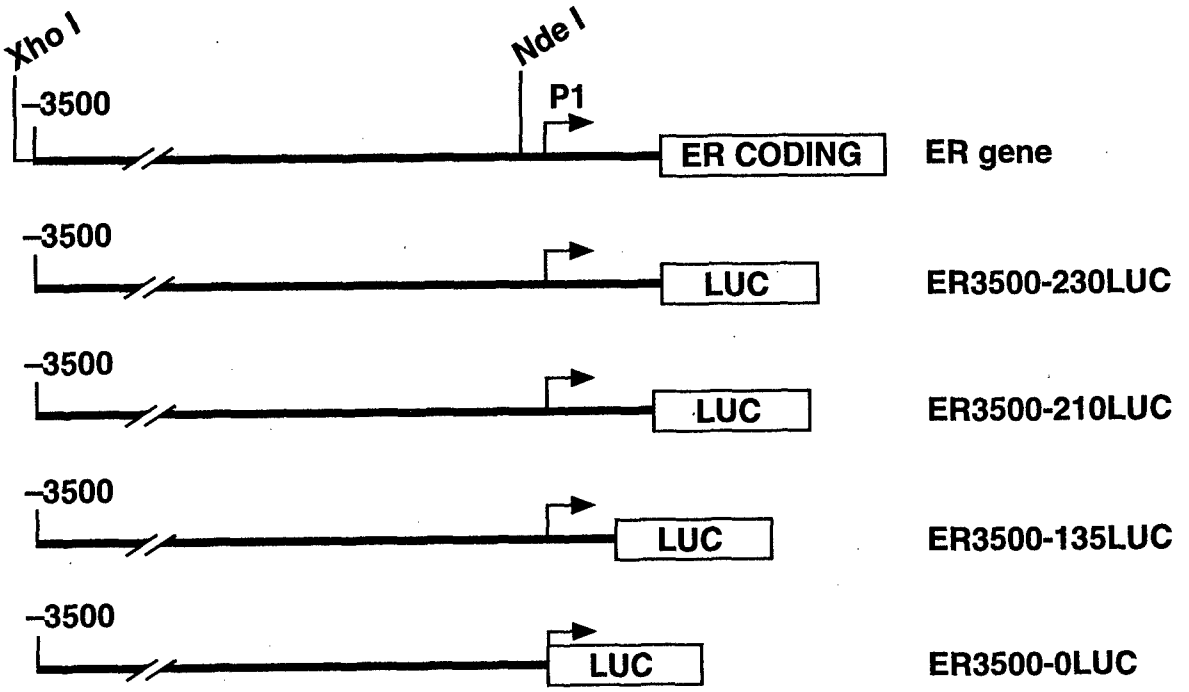

Xho I
Nde I
–3500
P1
ER CODING
ER gene
–3500
LUC
ER3500-230LUC
–3500
LUC
ER3500-210LUC
–3500
LUC
ER3500-135LUC
–3500
LUC
ER3500-0LUC

Figure 2. Luciferase activity from ER promoter constructs.

Plasmid DNA was transfected into T47D (ER-positive) or MDA-MB-231 (ER-negative) breast carcinoma cells and assayed for luciferase activity. Data presented is corrected for transfection efficiency. Data was normalized with the activity obtained from the largest plasmid in T47D as 100%. Standard error is shown by error bars. (A) Representative results of expression from 5' deletion constructs. Values shown are average of four separate transfection experiments. (B) Representative results of expression from 3' deletion constructs. Values shown are average of eight separate transfections. Data for ER3500-230p1d1 is average of two separate transfections performed in triplicate and is representative of other transfections performed.

Figure 2A.

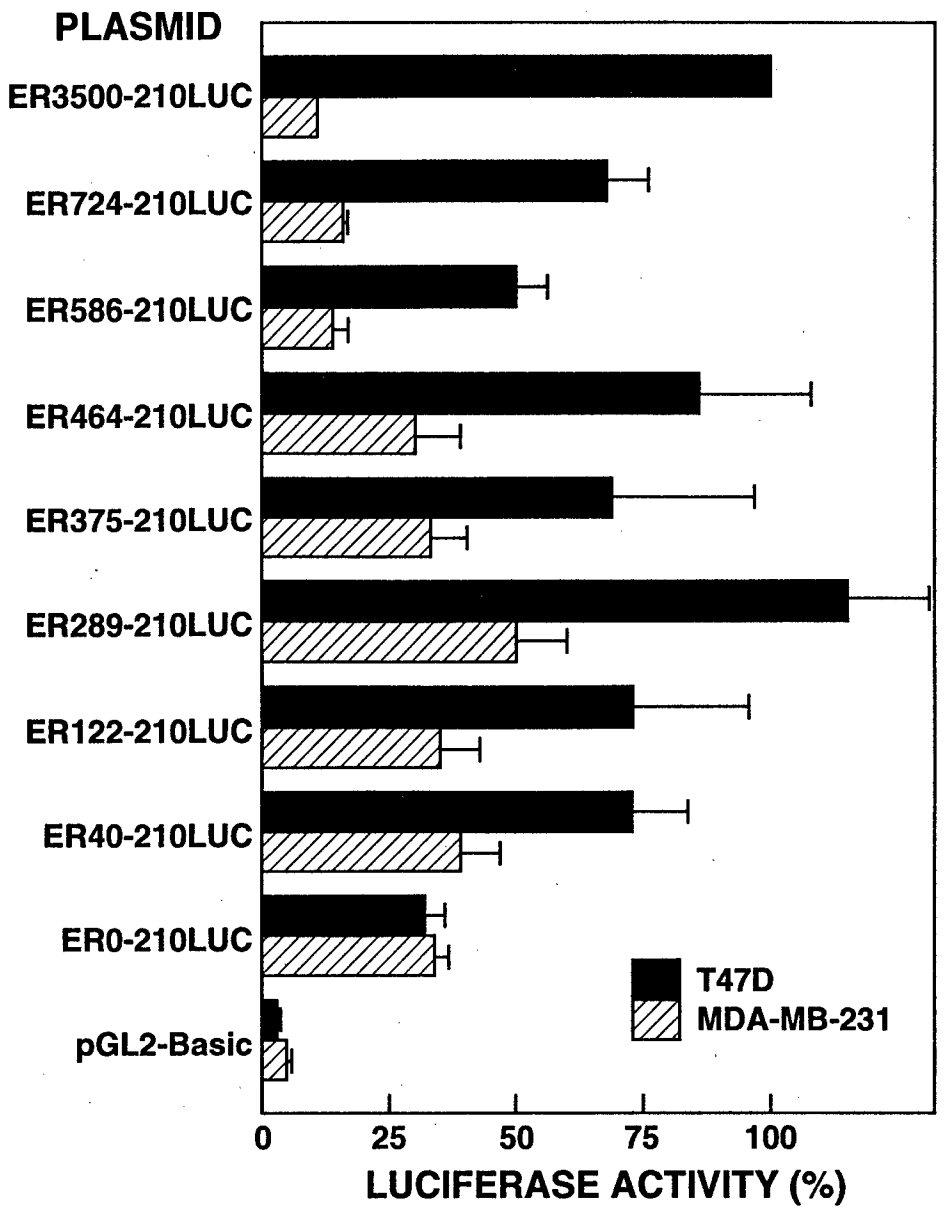

PLASMID
ER3500-210LUC
ER724-210LUC
ER586-210LUC
ER464-210LUC
ER375-210LUC
ER289-210LUC
ER122-210LUC
ER40-210LUC
ER0-210LUC
pGL2-Basic
T47D
MDA-MB-231
0
25
50
75
100
LUCIFERASE ACTIVITY (%)

Figure 2B.

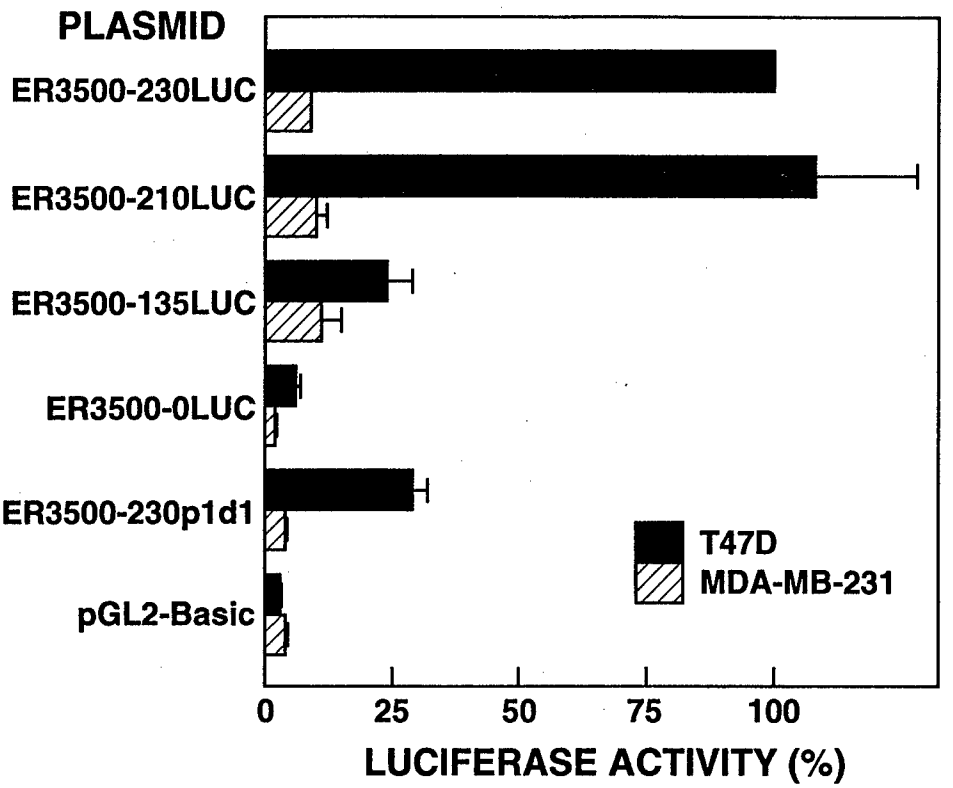

PLASMID
ER3500-230LUC
ER3500-210LUC
ER3500-135LUC
ER3500-0LUC
ER3500-230p1d1
pGL2-Basic
T47D
MDA-MB-231
0
25
50
75
100
LUCIFERASE ACTIVITY (%)

Figure 3. Gel shift assay using 80 bp probe.

Gel shift assay was performed using whole cell extracts from cells as shown. The probe is a radiolabelled 80 bp DNA fragment from ER promoter sequences +132 to +211. Gel shift competition involved addition of 500 fold molar excess cold competitor from sequences of ER gene as shown. In (-) lanes there is no competitor added.

Figure 3.

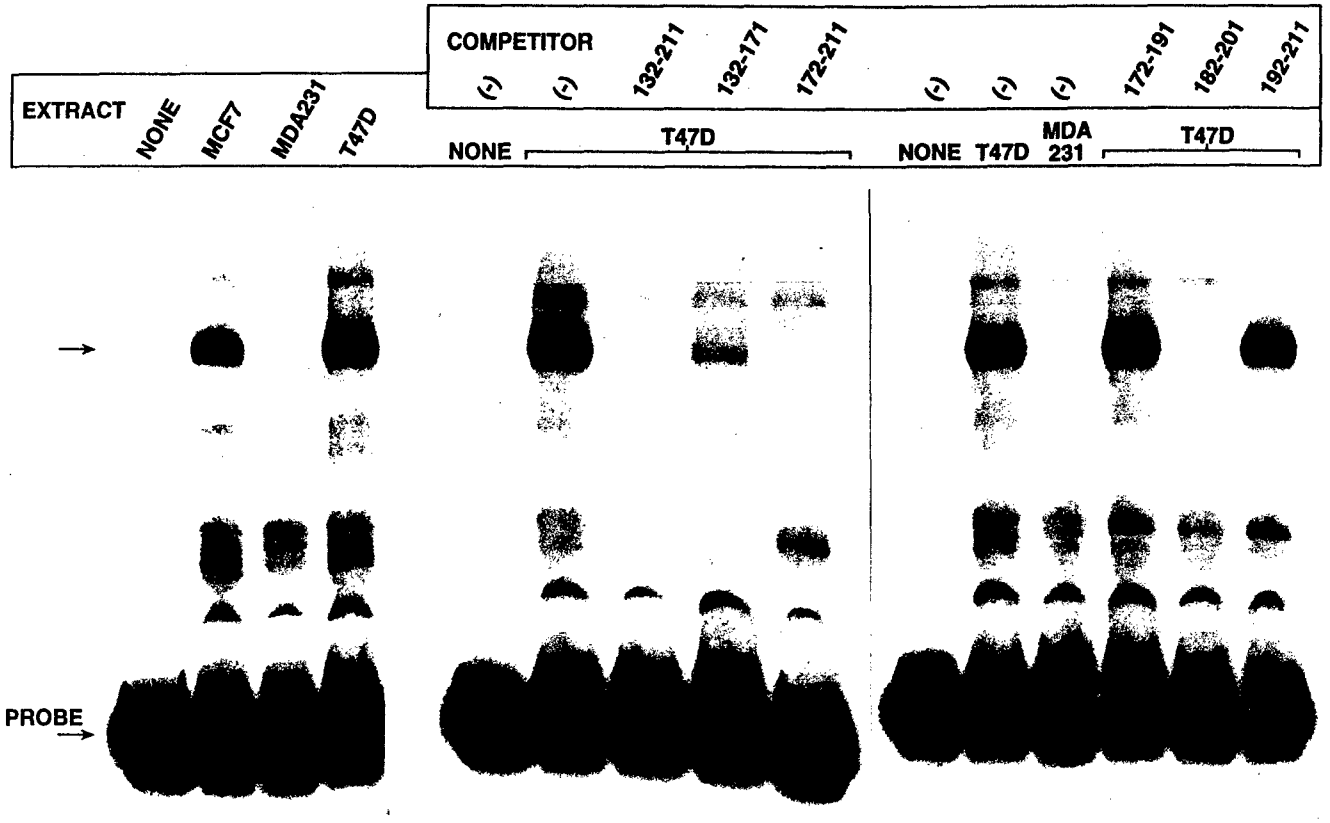

EXTRACT
NONE
MCF7
MDA231
T47D
COMPETITOR
(-)
(-)
132-211
132-171
172-211
(-)
(-)
(-)
172-191
182-201
192-211
NONE
T47D
NONE
T47D
MDA 231
T47D
PROBE

Figure 4. ERF-1 expression in human cell lines.

ERF-1 expression determined by gel-shift assay using whole cell extracts prepared from cell lines shown and reacted with a 72 bp probe encompassing proximal and distal ERF-1 sites. (A) Human breast carcinoma lines; ER-positive: MCF7, T47D, BT20 and ER-negative: MDA-MB-231 and HBL-100. HMEC are normal human mammary epithelial cells. (B) Other human cell lines; HeLa are cervical carcinoma cells; Daudi are lymphocytes; RL95-2, HEC 1B, HEC 1A, and ECC-1 are human endometrial carcinoma cell lines (see text for details).

Figure 4.

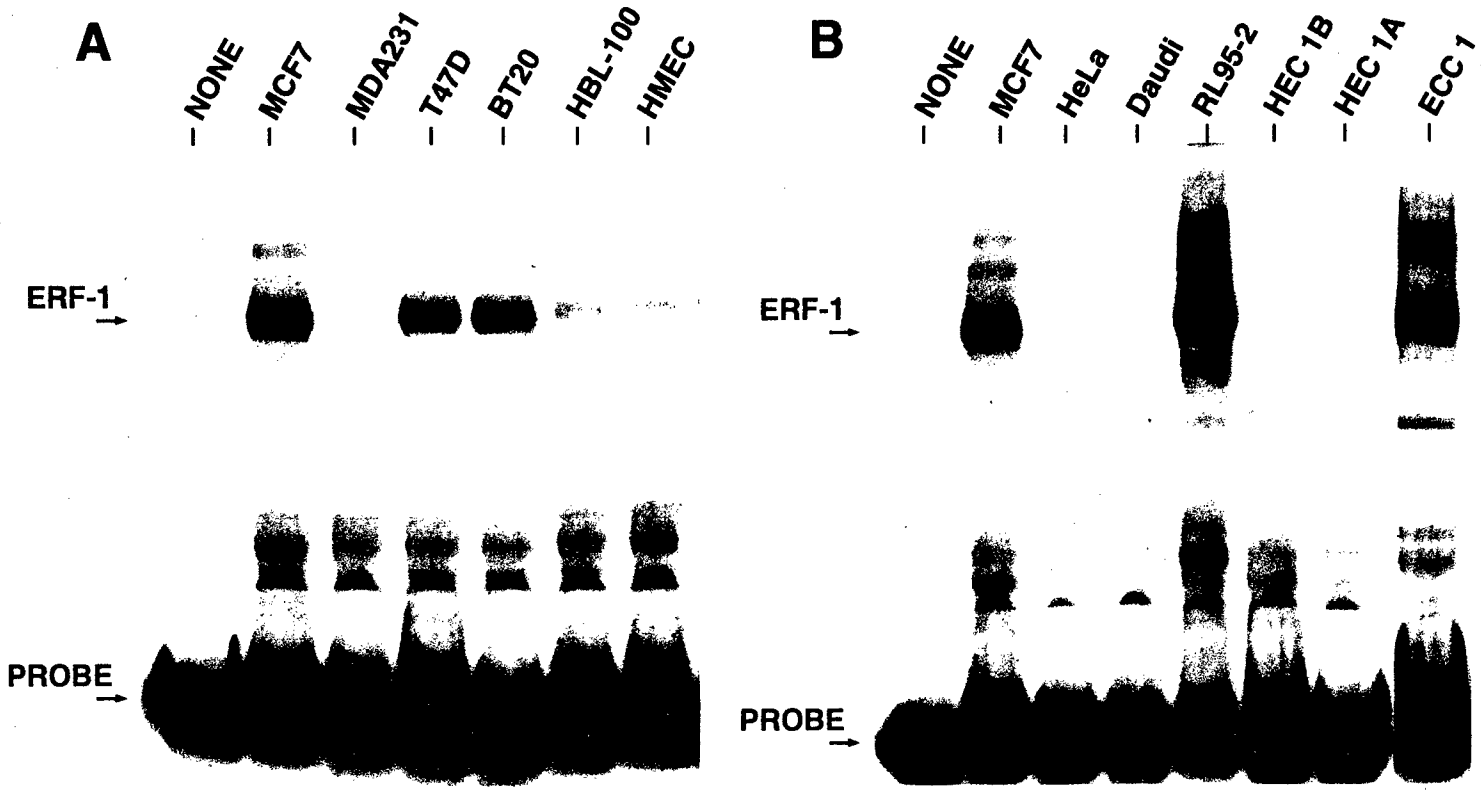

A
NONE
MCF7
MDA231
T47D
BT20
HBL-100
HMEC
ERF-1
PROBE
B
NONE
MCF7
HeLa
Daudi
RL95-2
HEC 1B
HEC 1A
ECC 1
ERF-1
PROBE

Figure 5. Mapping ERF-1 binding sites.

Gel shift competitions were performed using whole cell extract from cells shown. Probe used in both A and B is a 72 bp DNA fragment from ER promoter sequences +130 to +201. (A) Competitions using mutant distal binding sites. Sequence of wild-type distal site (dwt) and position of mutated sequence is shown for each mutant sequence d1 - d5. (B) Proximal binding site defined using proximal mutants p1 - p4. Below gel is shown sequences of distal site (dwt) and proximal site wild-type sequence aligned to highlight homology. Sequence of proximal site wild-type sequence (pwt) and four proximal site mutants p1 - p4 is shown.

Figure 5.

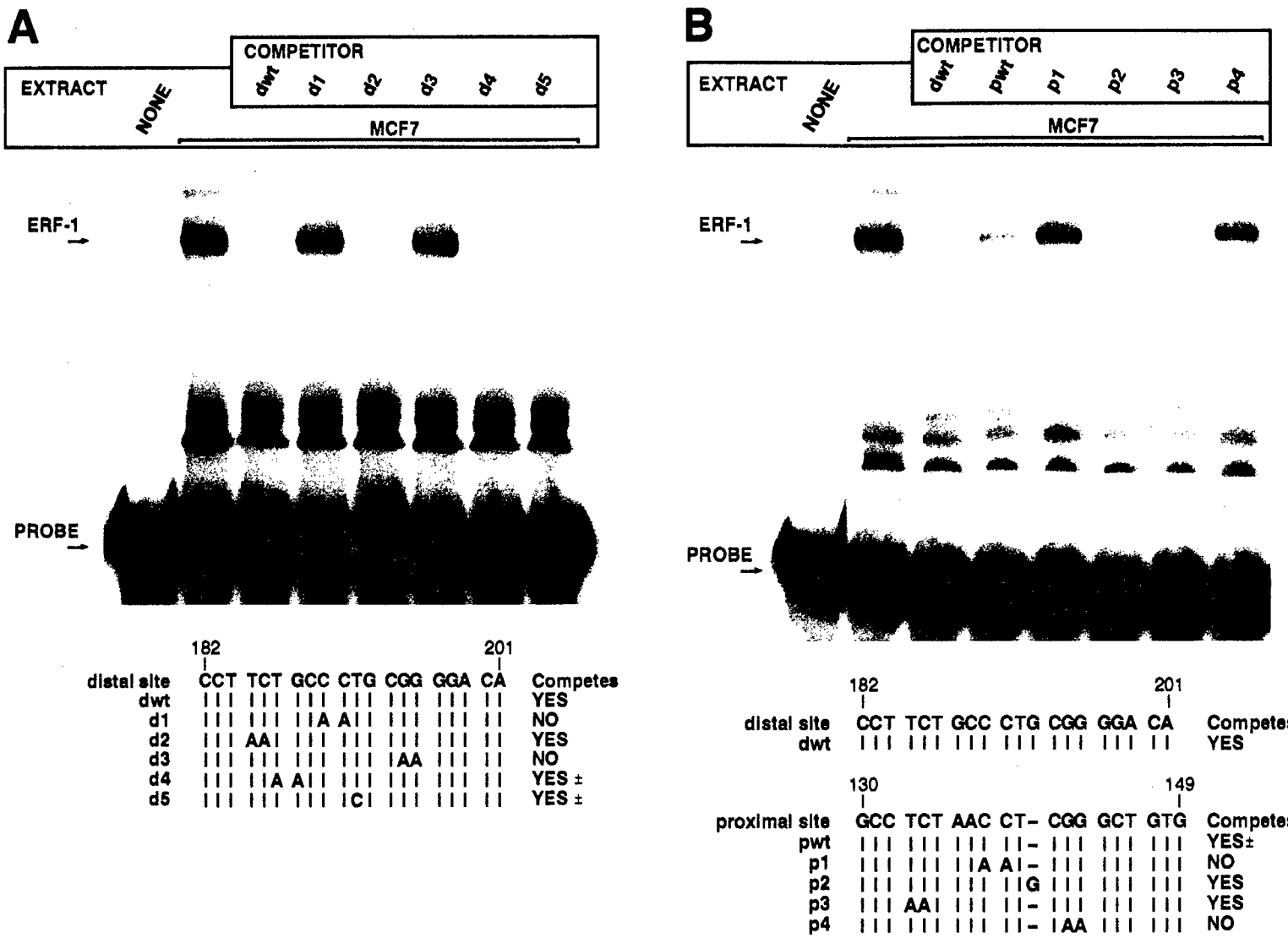

A
EXTRACT
NONE
COMPETITOR
dwt
d1
d2
d3
d4
d5
MCF7
ERF-1
PROBE
182
201
distal site CCT TCT GCC CTG CGG GGA CA Competes
dwt YES
d1 NO
d2 YES
d3 NO
d4 YES ±
d5 YES ±
B
EXTRACT
NONE
COMPETITOR
dwt
pwt
p1
p2
p3
p4
MCF7
ERF-1
PROBE
182
201
distal site CCT TCT GCC CTG CGG GGA CA Competes
dwt YES
130
149
proximal site GCC TCT AAC CT- CGG GCT GTG Competes
pwt YES±
p1 NO
p2 YES
p3 YES
p4 NO

Figure 6. Mutational Analysis of ERF-1 Binding Site

Gel shift analysis using nuclear extract prepared from MCF7 cells (lanes 2-19) or MDA-MB-231 (lane 20) with 30 bp wild-type oligonucleotide as probe. Competition with mutant oligonucleotides as listed.

Figure 6.

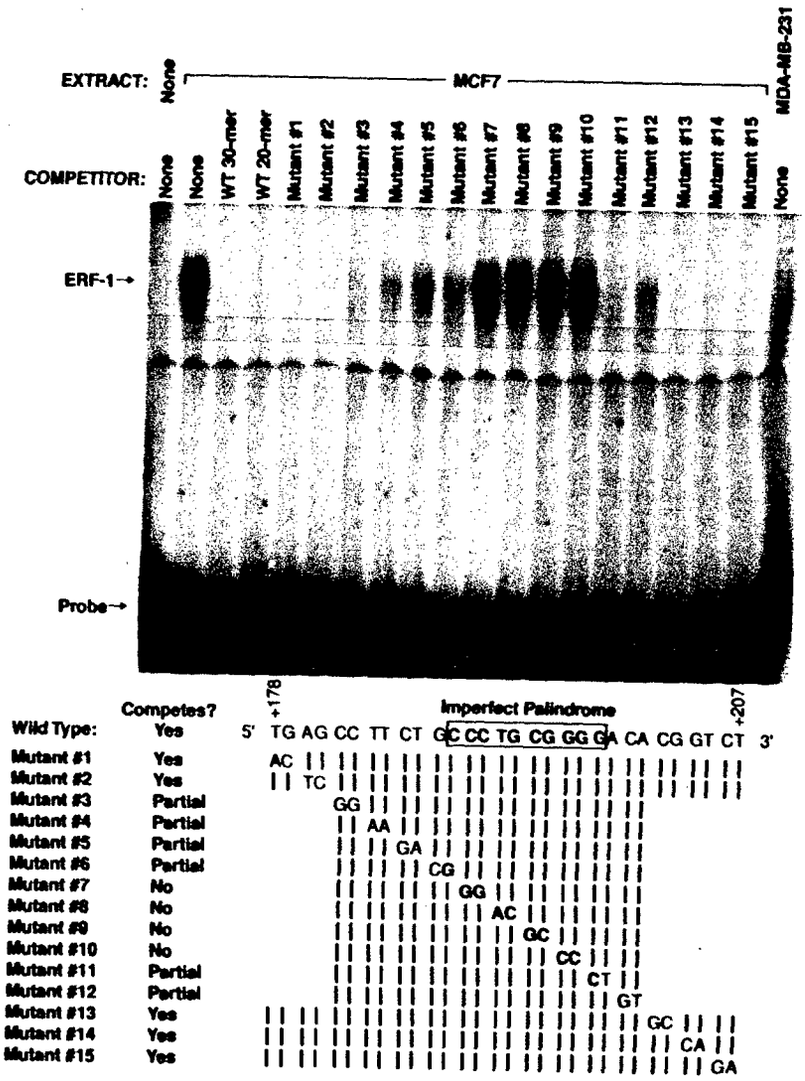

EXTRACT:
None
MCF7
MDA-MB-231
COMPETITOR:
None
None
WT 30-mer
WT 20-mer
Mutant #1
Mutant #2
Mutant #3
Mutant #4
Mutant #5
Mutant #6
Mutant #7
Mutant #8
Mutant #9
Mutant #10
Mutant #11
Mutant #12
Mutant #13
Mutant #14
Mutant #15
None
ERF-1→
Probe→
Competes?
+178
Imperfect Palindrome
+207
Wild Type: Yes 5' TG AG CC TT CT GC CC TG CG GG GA CA CG GT CT 3'
Mutant #1 Yes AC
Mutant #2 Yes TC
Mutant #3 Partial GG
Mutant #4 Partial AA
Mutant #5 Partial GA
Mutant #6 Partial CG
Mutant #7 No GG
Mutant #8 No AC
Mutant #9 No GC
Mutant #10 No CC
Mutant #11 Partial CT
Mutant #12 Partial GT
Mutant #13 Yes GC
Mutant #14 Yes CA
Mutant #15 Yes GA

Figure 7. Purification of ERF-1.

Protocol for purification of ERF-1 protein using ion exchange and DNA affinity chromatography. See text for details.

Figure 7.

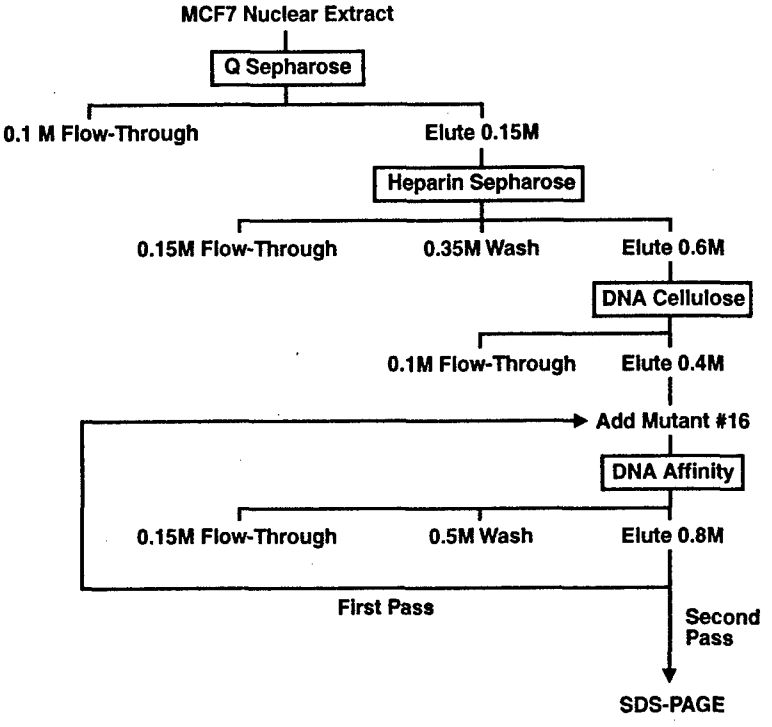

MCF7 Nuclear Extract
Q Sepharose
0.1 M Flow-Through
Elute 0.15M
Heparin Sepharose
0.15M Flow-Through
0.35M Wash
Elute 0.6M
DNA Cellulose
0.1M Flow-Through
Elute 0.4M
Add Mutant #16
DNA Affinity
0.15M Flow-Through
0.5M Wash
Elute 0.8M
First Pass
Second Pass
SDS-PAGE

Figure 8. Size Determination of ERF-1

A: ERF-1 complex was resolved by gel-shift analysis and the gel was subjected to UV-crosslinking. The bound complex was excised and resolved on SDS-PAGE as shown. The location of molecular weight markers (kD) is shown to the left of the gel lane. B: Purified extract was electrophoresed on 8% SDS-PAGE and sectioned into 12 gel slices. The approximate location of gel slices 1 through 12 are shown to the right of the gel lane. Molecular weight markers (kD) are listed to the left of the lane labelled markers. C: Gel shift analysis of proteins renatured from gel slices in B. Fraction number 7 demonstrated ERF-1 activity.

Figure 8.

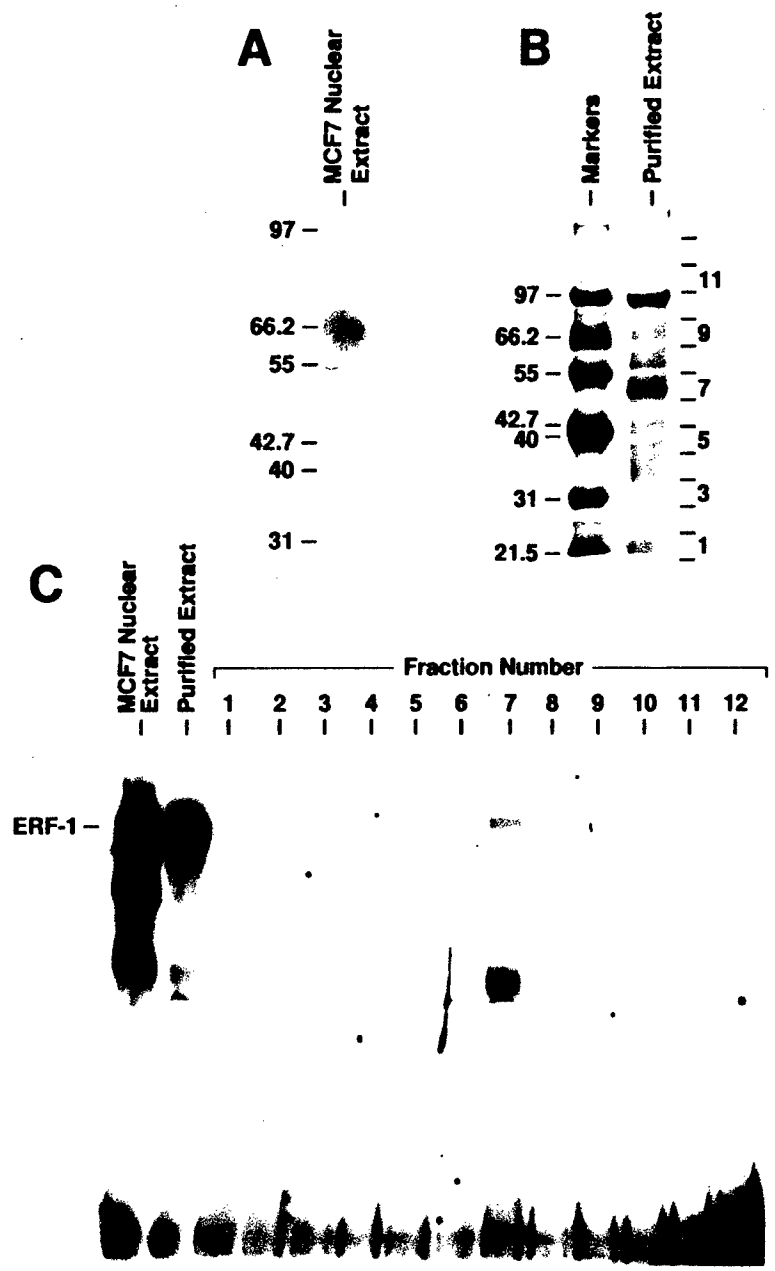
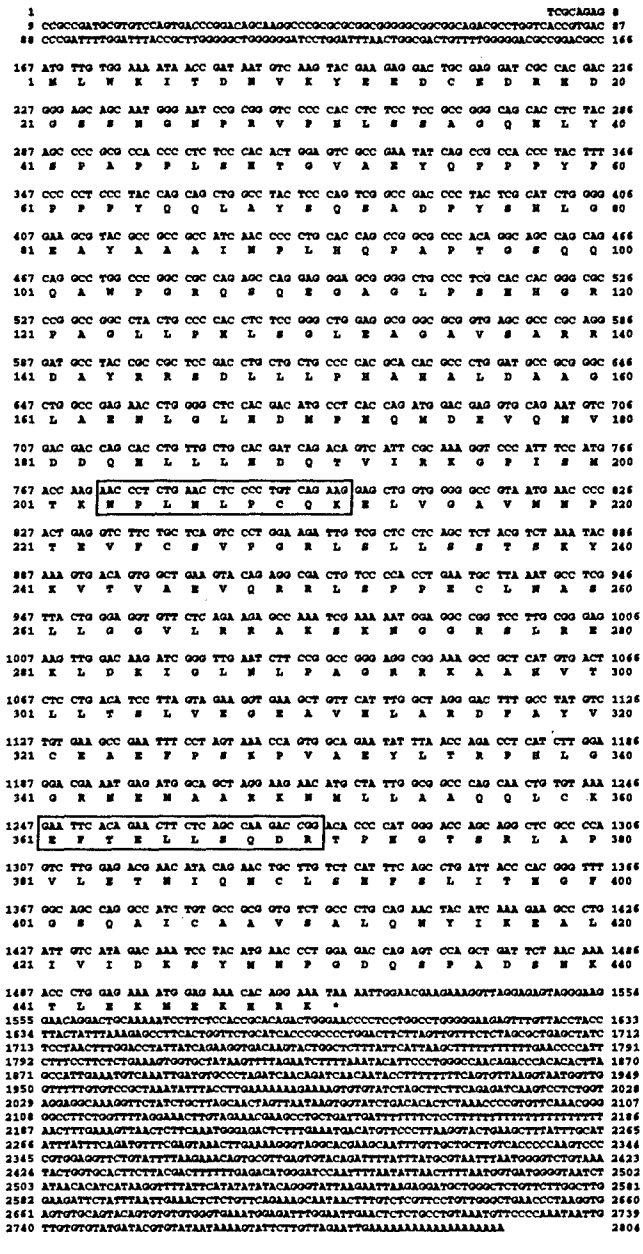

A
MCF7 Nuclear Extract
97
66.2
55
42.7
40
31
B
Markers
Purified Extract
97
66.2
55
42.7
40
31
21.5
11
9
7
5
3
1
C
MCF7 Nuclear Extract
Purified Extract
Fraction Number
1 2 3 4 5 6 7 8 9 10 11 12
ERF-1

Figure 9. ERF-1 cDNA

A: Complete sequence of ERF-1 cDNA with predicted amino acid sequence. Boxes indicate peptide sequences obtained from protein sequencing of purified ERF-1. B: Comparison of ERF-1 protein sequence with AP2α protein aligned to highlight similarities. (*) indicates identical amino acids; (.) indicates conserved substitution.

Figure 9.

A.



B.

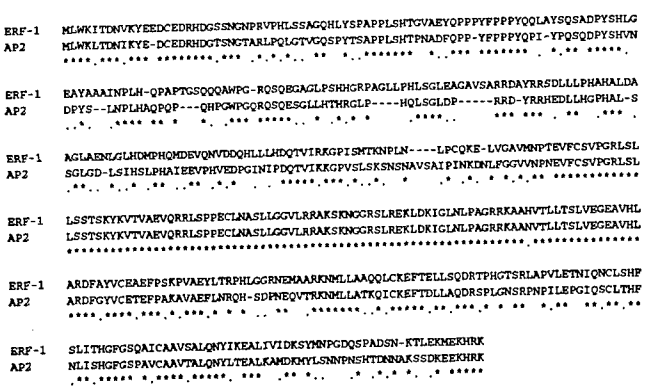


Figure 10. *In vitro* synthesized ERF-1

A: *In vitro* transcription/translation from the ERF-1 cDNA using T3 polymerase (sense) (lane 1) or T7 polymerase (antisense) (lane 3). *In vitro* synthesized ERF-1 was immunoprecipitated using AP2 polyclonal antisera (lane 2).

B: Gel shift analysis of *in vitro* synthesized ERF-1 (lanes 2-13) with mutant competitors compared to MCF7 nuclear extract (lane 1). *In vitro* ERF-1 complex supershifted with AP2 antisera (lane 13).

Figure 10.

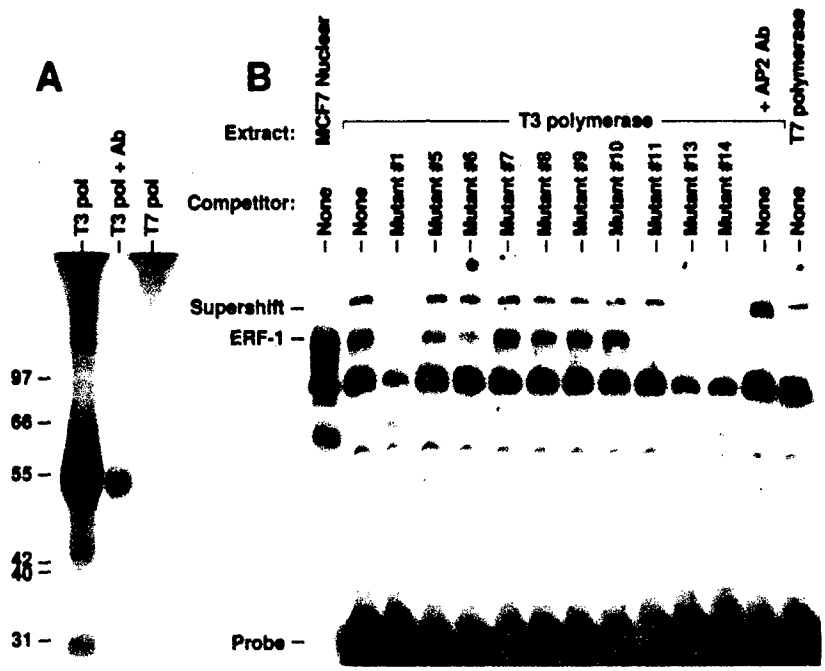

A
T3 pol
T3 pol + Ab
T7 pol
97 –
66 –
55 –
43 –
31 –
B
Extract:
MCF7 Nuclear
T3 polymerase
+ AP2 Ab
T7 polymerase
Competitor:
None
None
Mutant #1
Mutant #5
Mutant #6
Mutant #7
Mutant #8
Mutant #9
Mutant #10
Mutant #11
Mutant #13
Mutant #14
None
None
Supershift –
ERF-1 –
Probe –

Figure 11. Native ERF-1 complex supershift

MCF7 nuclear extract analyzed by gel shift with AP2 antisera (lanes 2-14) with competitors as shown compared to native ERF-1 complex (lane 1).

Figure 11.

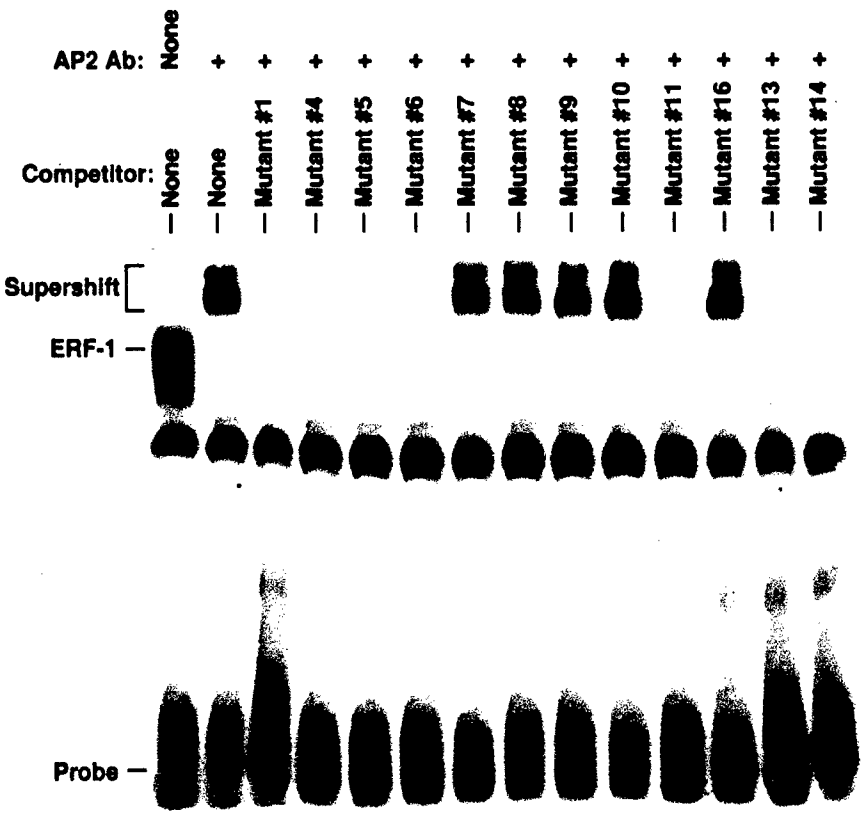

AP2 Ab:
None
+
+
+
+
+
+
+
+
+
+
+
+
Competitor:
— None
— None
— Mutant #1
— Mutant #4
— Mutant #5
— Mutant #6
— Mutant #7
— Mutant #8
— Mutant #9
— Mutant #10
— Mutant #11
— Mutant #16
— Mutant #13
— Mutant #14
Supershift
ERF-1 —
Probe —

Figure 12. Competitive gel shifts using AP2α with a panel of competitors

Affinity purified AP2α analyzed by gel shift with competitors as shown. OPT (GCCTGAGGG) labeled with [α -^{32}P] dCTP was used as a probe and was competed with increasing amounts of unlabeled competitor. Relative binding affinity compared to OPT is indicated next to gel shift competitions for each mutant competitor.

Figure 12.

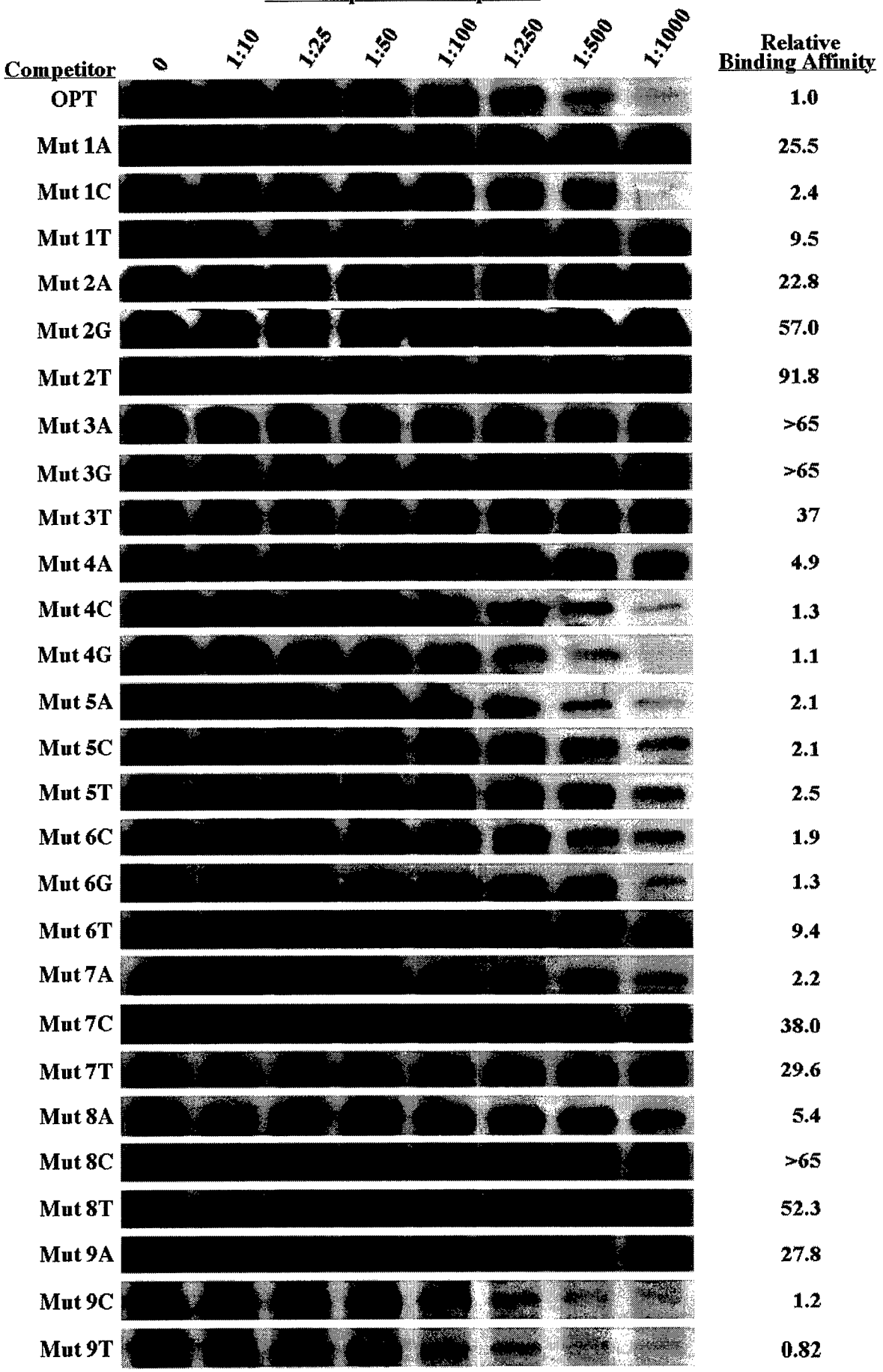

Ratio of probe to competitor
Competitor
0
1:10
1:25
1:50
1:100
1:250
1:500
1:1000
Relative Binding Affinity
OPT 1.0
Mut 1A 25.5
Mut 1C 2.4
Mut 1T 9.5
Mut 2A 22.8
Mut 2G 57.0
Mut 2T 91.8
Mut 3A >65
Mut 3G >65
Mut 3T 37
Mut 4A 4.9
Mut 4C 1.3
Mut 4G 1.1
Mut 5A 2.1
Mut 5C 2.1
Mut 5T 2.5
Mut 6C 1.9
Mut 6G 1.3
Mut 6T 9.4
Mut 7A 2.2
Mut 7C 38.0
Mut 7T 29.6
Mut 8A 5.4
Mut 8C >65
Mut 8T 52.3
Mut 9A 27.8
Mut 9C 1.2
Mut 9T 0.82

Figure 13. Competitive gel shifts using ERF-1 with a panel of competitors

ERF-1/GST fusion analyzed by gel shift with competitors as shown. OPT (GCCTGAGGG) labeled with [α -^{32}P] dCTP was used as a probe and was competed with increasing amounts of each unlabeled competitor. Relative binding affinity compared to OPT is indicated next to gel shift competitions for each mutant competitor.

Figure 13.

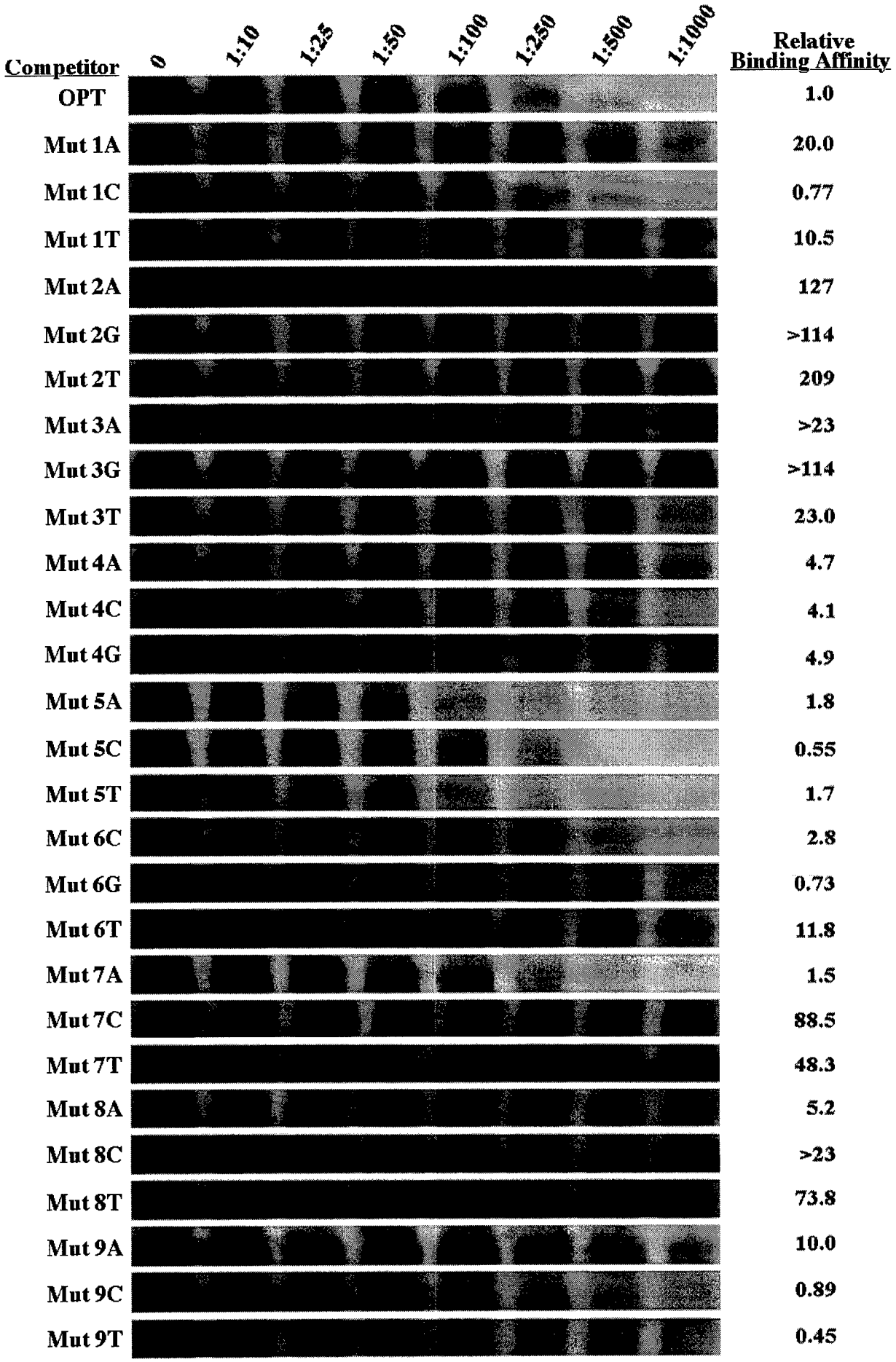
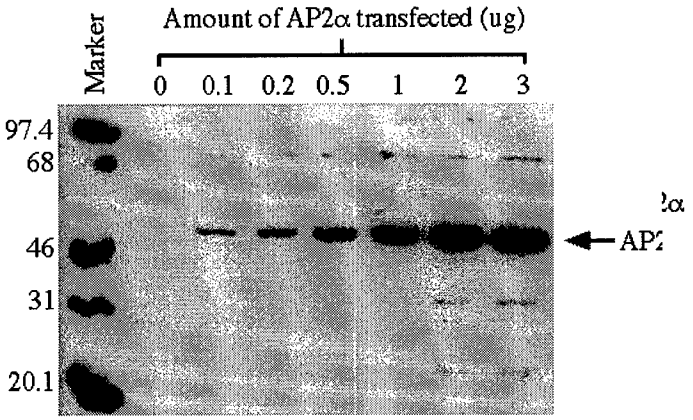

Ratio of probe to competitor
0
1:10
1:25
1:50
1:100
1:250
1:500
1:1000
Competitor
Relative Binding Affinity
OPT 1.0
Mut 1A 20.0
Mut 1C 0.77
Mut 1T 10.5
Mut 2A 127
Mut 2G >114
Mut 2T 209
Mut 3A >23
Mut 3G >114
Mut 3T 23.0
Mut 4A 4.7
Mut 4C 4.1
Mut 4G 4.9
Mut 5A 1.8
Mut 5C 0.55
Mut 5T 1.7
Mut 6C 2.8
Mut 6G 0.73
Mut 6T 11.8
Mut 7A 1.5
Mut 7C 88.5
Mut 7T 48.3
Mut 8A 5.2
Mut 8C >23
Mut 8T 73.8
Mut 9A 10.0
Mut 9C 0.89
Mut 9T 0.45

Figure 14. *In vivo* AP2α activation of OPT/IL-2 *vs* M3T/IL-2 reporter constructs

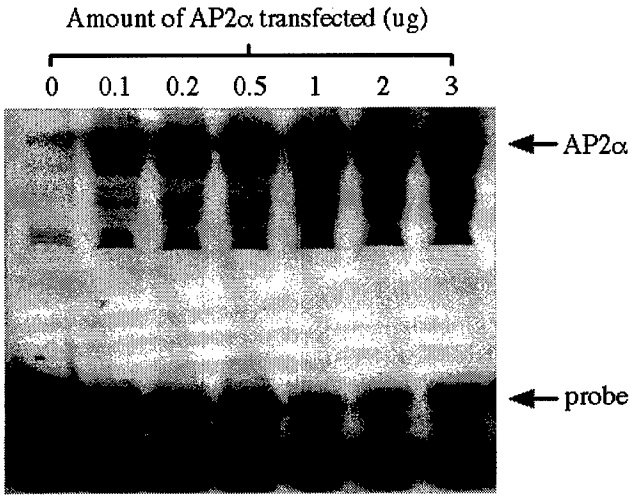
A. Western blot analysis of whole cell extracts from transfection of COS cells with AP2α and AP2 binding site reporter constructs. IL-2 reporter construct contains only the minimal IL-2 promoter and is a measure of background activity in the cell. OPT/IL-2 contains the optimal AP2 binding site (GCCTGAGGG) while M3T/IL-2 contains a mutated AP2 binding site (GCTTGAGGG). B. Gel shift assay of whole cell extracts as described above. C. Luciferase assay of cell extracts from transfections described above. Corrected Luciferase units determined as described in text.

Figure 14.

A.



B.



C.

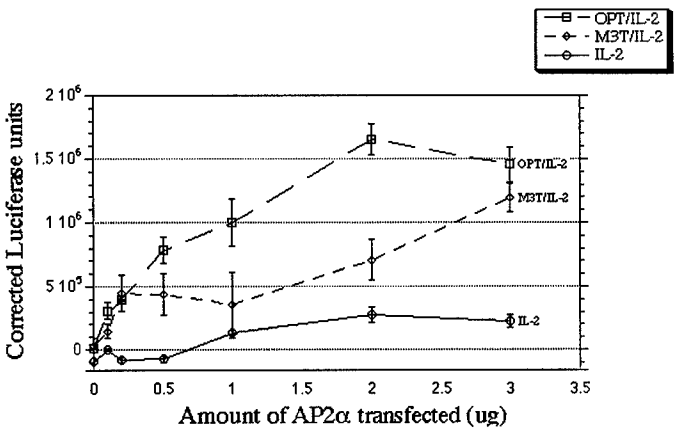
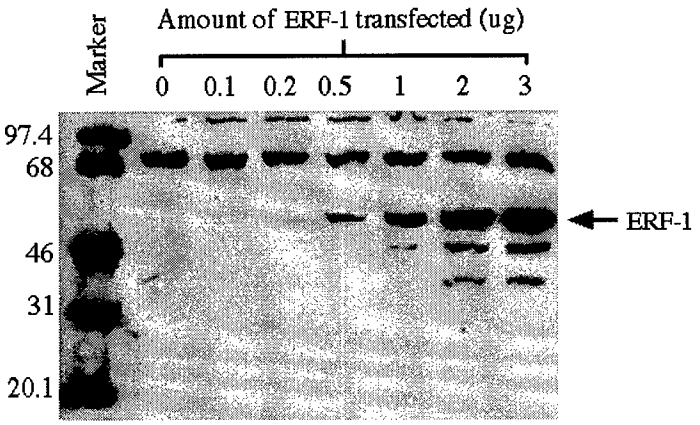


Figure 15. *In vivo* ERF-1 activation of OPT/IL-2 *vs* M3T/IL-2 reporter constructs

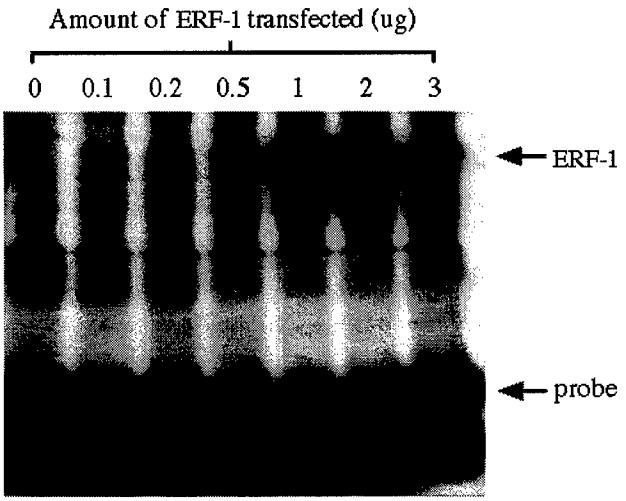
A. Western blot analysis of whole cell extracts from transfection of COS cells with ERF-1 and AP2 binding site reporter constructs. IL-2 reporter construct contains only the minimal IL-2 promoter and is a measure of background activity in the cell. OPT/IL-2 contains the optimal AP2 binding site (GCCTGAGGG) while M3T/IL-2 contains a mutated AP2 binding site (GCTTGAGGG). B. Gel shift assay of whole cell extracts as described above. C. Luciferase assay of cell extracts from transfections described above. Corrected Luciferase units determined as described in text.

Figure 15.

A.



B.



C.

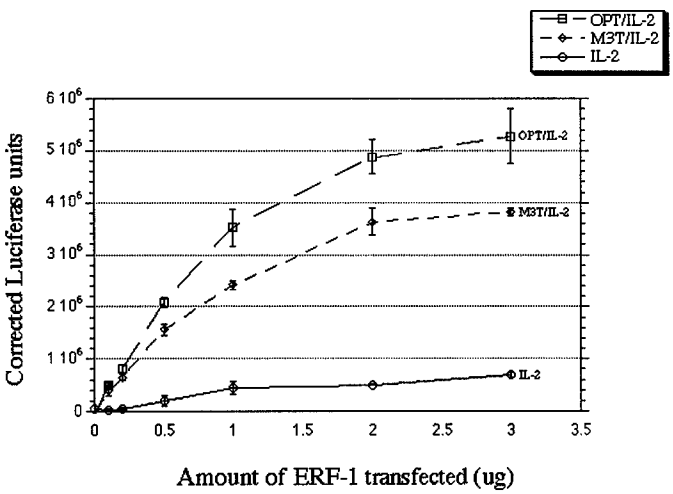


Figure 16. *In vivo* activation of ER3500-230LUC by AP2α and ERF-1

A. Luciferase assay of cell extracts from transfection of HepG2 cells with AP2α and ERF-1 and ER3500-230LUC or pGL2-BASIC. B. Luciferase assay of cell extracts from transfection of MDA-MB-231 cells with AP2α and ERF-1 and ER3500-230LUC or pGL2-BASIC. Corrected Luciferase units determined as described in text.

Figure 16.

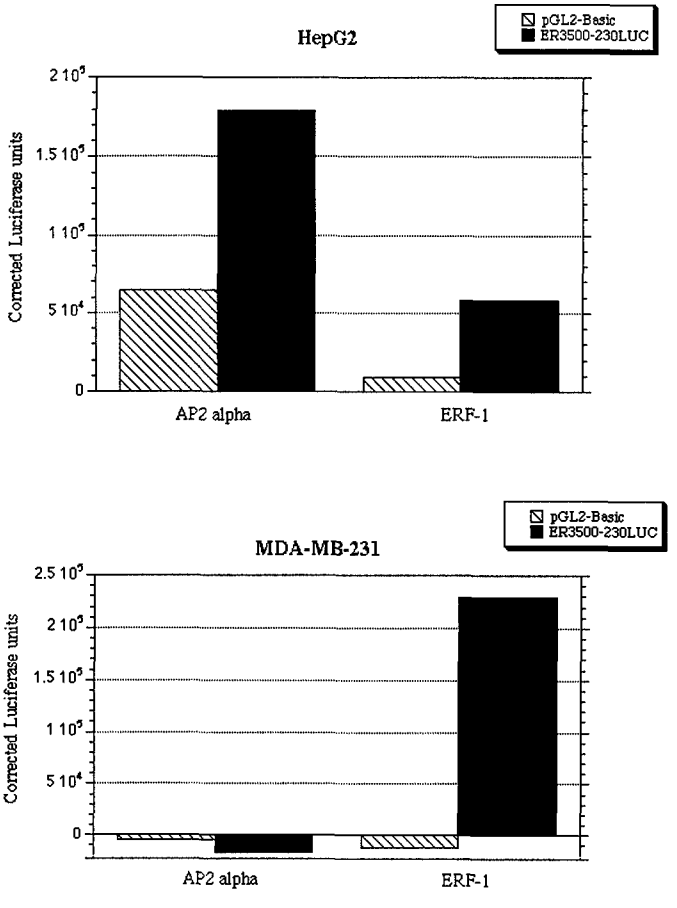

HepG2
pGL2-Basic
ER3500-230LUC
Corrected Luciferase units
2 10⁵
1.5 10⁵
1 10⁵
5 10⁴
0
AP2 alpha
ERF-1
MDA-MB-231
pGL2-Basic
ER3500-230LUC
Corrected Luciferase units
2.5 10⁵
2 10⁵
1.5 10⁵
1 10⁵
5 10⁴
0
AP2 alpha
ERF-1

Table 1. Summary of all AP2α binding sites

Position	1	2	3	4	5	6	7	8	9
G	185	0	0	46	61	67	91	137	79
A	0	0	0	22	19	55	53	19	9
T	0	0	0	46	48	19	11	13	19
C	0	185	185	71	57	44	30	16	78
G	100%	0%	0%	25%	33%	36%	49%	74%	43%
A	0%	0%	0%	12%	10%	30%	29%	10%	5%
T	0%	0%	0%	25%	26%	10%	6%	7%	10%
C	0%	100%	100%	38%	31%	24%	16%	9%	42%
	G	C	C	C/G/T	G/C/T	G/A/C	G/A	G	G/C

Number of each nucleotide found in each position listed in top half of table. Percent of each nucleotide at each position listed in bottom half.

Table 2. Summary of all ERF-1 binding sites

Position	1	2	3	4	5	6	7	8	9
G	146	0	0	31	58	57	93	116	63
A	0	0	0	4	14	46	28	12	11
T	0	0	0	46	28	8	12	9	16
C	0	146	146	65	46	35	13	9	56
G	100%	0%	0%	21%	40%	39%	64%	80%	43%
A	0%	0%	0%	3%	10%	32%	19%	8%	8%
T	0%	0%	0%	31%	19%	5%	8%	6%	11%
C	0%	100%	100%	45%	31%	24%	9%	6%	38%
	G	C	C	C/t/g	G/C	G/a/c	G	G	G/c

Number of each nucleotide found in each position listed in top half of table. Percent of each nucleotide at each position listed in bottom half.

Table 3. Summary of unique AP2α binding sites

Position	1	2	3	4	5	6	7	8	9
G	46	0	0	20	21	20	25	40	26
A	0	0	0	1	2	18	17	4	2
T	0	0	0	15	7	4	0	1	5
C	0	46	46	10	16	4	4	1	13
G	100%	0%	0%	43%	46%	43%	54%	87%	57%
A	0%	0%	0%	2%	4%	39%	37%	9%	4%
T	0%	0%	0%	33%	15%	9%	0%	2%	11%
C	0%	100%	100%	22%	35%	9%	9%	2%	28%
	G	C	C	G/T/C	G/C	G/A	G/A	G	G/C

Only clones containing a single binding site are listed.

Table 4. Summary of unique ERF-1 binding sites

Position	1	2	3	4	5	6	7	8	9
G	48	0	0	12	20	20	29	40	25
A	0	0	0	0	2	15	15	6	5
T	0	0	0	13	6	3	2	2	7
C	0	48	48	23	20	10	2	0	11
G	100%	0%	0%	25%	42%	42%	60%	83%	52%
A	0%	0%	0%	0%	4%	31%	31%	13%	10%
T	0%	0%	0%	27%	13%	6%	4%	4%	15%
C	0%	100%	100%	48%	42%	21%	4%	0%	23%
	G	C	C	C/t/g	G/C	G/a/c	G	G	G/c

Only clones containing a single binding site are listed.

Table 5. Summary of AP2 alpha clones from revised binding site selection

Position	1	2	3	4	5	6	7	8	9
G	19	0	0	8	18	6	27	27	11
A	0	0	0	1	1	9	0	0	0
T	0	0	0	8	2	6	0	0	0
C	8	27	27	10	6	6	0	0	16
G	70%	0%	0%	30%	67%	22%	100%	100%	41%
A	0%	0%	0%	4%	4%	33%	0%	0%	0%
T	0%	0%	0%	30%	7%	22%	0%	0%	0%
C	30%	100%	100%	37%	22%	22%	0%	0%	59%
	G	C	C	C/T/G	G	N	G	G	C/G

Number of each nucleotide found in each position listed in top half of table. Percent of each nucleotide at each position listed in bottom half.

Table 6. Summary of ERF-1 clones from revised binding site selection

Position	1	2	3	4	5	6	7	8	9
G	21	0	0	5	14	7	25	25	13
A	0	0	0	2	4	10	0	0	0
T	0	0	0	12	3	2	0	0	0
C	4	25	25	6	4	6	0	0	12
G	84%	0%	0%	20%	56%	28%	100%	100%	52%
A	0%	0%	0%	8%	16%	40%	0%	0%	0%
T	0%	0%	0%	48%	12%	8%	0%	0%	0%
C	16%	100%	100%	24%	16%	24%	0%	0%	48%
	G	C	C	T/c/g	G	A/g/c	G	G	G/C

Number of each nucleotide found in each position listed in top half of table. Percent of each nucleotide at each position listed in bottom half.

BIBLIOGRAPHY

1. de Coninck, E.C., McPherson, L.A. & Weigel, R.J. Transcriptional Regulation of Estrogen Receptor in Breast Carcinoma. *Mol. Cell. Biol.* **15**, 2191-2196 (1995).

2. McPherson, L.A., Baichwal, V.R. & Weigel, R.J. Identification of ERF-1 as a member of the AP2 transcription factor family. *Proc. Natl. Acad. Sci. USA* **94**, 4342-4347 (1997).

LIST OF PERSONNEL

1. Ronald J Weigel, P.I.

2. Ellen C. De Coninck, Professional-nonstudent

3. Renee V. Hoch, Professional-nonstudent

4. Wayne W. Kuang, Research Asst.-Grad Student

5. Malavika G. Ghosh, Research Asst.-Grad Student

6. Lisa A. McPherson, Student Research Asst.- Postdoctoral fellow

7. Devon A. Thompson, Student Research Asst.- Postdoctoral fellow

8. Eric R. Schuur, Technical-Nonstudent

9. Andrew J. Krystkowiak, Technical-Nonstudent

Rec'd 1/11/2000

DEPARTMENT OF THE ARMY
US ARMY MEDICAL RESEARCH AND MATERIEL COMMAND
504 SCOTT STREET
FORT DETRICK, MARYLAND 21702-5012

REPLY TO
ATTENTION OF:

MCMR-RMI-S (70-1y) 4 Jan 00

MEMORANDUM FOR Administrator, Defense Technical Information Center, ATTN: DTIC-OCA, 8725 John J. Kingman Road, Fort Belvoir, VA 22060-6218

SUBJECT: Request Change in Distribution Statement

1. The U.S. Army Medical Research and Materiel Command has reexamined the need for the limitation assigned to technical reports written for the attached Grants. Request the limited distribution statements for Accession Document Numbers listed be changed to "Approved for public release; distribution unlimited." This report should be released to the National Technical Information Service.

2. Point of contact for this request is Ms. Judy Pawlus at DSN 343-7322 or by email at Judy.Pawlus@amedd.army.mil.

FOR THE COMMANDER:

PHYLIS M. RINEHART
Deputy Chief of Staff for
Information Management

94-J-4353 AD-B246 269

Completed 1-14-00 cr